

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 19, 2003, 16:55:48 ; Search time 2430.79 Seconds
(without alignments)
10804.703 Million cell updates/sec
Title: US-09-899-303A-3
Perfect score: 642
Sequence: 1 ATGCCCGGTCTTCTTC.....TACTCTTGCTCTCTAATAG 642

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sv.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	642	100.0	642	6	A48665	A48665 Sequence 3
2	642	100.0	642	6	AR157324	AR157324 Sequence
3	642	100.0	642	6	AX452752	AX452752 Sequence
4	642	100.0	642	6	AX685004	AX685004 Sequence
5	628.2	97.9	795	6	A48667	A48667 Sequence 5
6	628.2	97.9	795	6	AR157325	AR157325 Sequence
7	628.2	97.9	795	6	AX452754	AX452754 Sequence
8	628.2	97.9	795	6	AX685006	AX685006 Sequence
9	619.8	96.5	2082	6	A48709	A48709 Sequence 47
10	619.8	96.5	2082	6	AR157350	AR157350 Sequence
11	619.8	96.5	2082	6	AX452796	AX452796 Sequence
12	619.8	96.5	2082	6	AX685048	AX685048 Sequence
13	619.8	96.5	2433	6	A48711	A48711 Sequence 49
14	619.8	96.5	2433	6	AR157351	AR157351 Sequence
15	619.8	96.5	2433	6	AX452798	AX452798 Sequence
16	619.8	96.5	2433	6	AX685050	AX685050 Sequence
17	575	89.6	9379	14	AF207766	AF207766 Hepatitis
18	573.4	89.3	9410	14	HPC1K12	D50481 Hepatitis C
19	569.6	88.7	9418	14	HCV132996	AJ132996 Hepatitis
20	568.6	88.6	9386	14	AF165056	AF165056 Hepatitis
21	567	88.3	9386	14	AF165055	AF165055 Hepatitis
22	567	88.3	9410	14	HPC1K12	D50485 Hepatitis C
23	567	88.3	9598	14	AB049101	AB049101 Hepatitis
24	566.4	88.2	3461	6	E07544	E07544 Gene fragment
25	566.4	88.2	3461	6	E09288	E09288 cdNA encodi
26	565.4	88.1	1880	14	HPCSTRJ4	D00832 Hepatitis C
27	565.4	88.1	2540	6	E04260	E04260 cdNA encodi
28	565.4	88.1	2540	6	E04805	E04805 cdNA to 5'
29	565.4	88.1	2540	6	E07391	E07391 cdNA encodi
30	565.4	88.1	3296	14	AB008441	AB008441 Hepatitis
31	565.4	88.1	9033	14	HCV238800	AJ238800 Hepatitis
32	565.4	88.1	9448	14	HPCJ483	D13558 Hepatitis C
33	565.4	88.1	9456	14	HPCBNA	D10934 Hepatitis C
34	565.4	88.1	9605	6	AX739971	AX739971 Sequence
35	565.4	88.1	9605	14	HCV238799	AJ238799 Hepatitis
36	565.4	88.1	11076	6	AX036252	AX036252 Sequence
37	565.4	88.1	11076	6	AX036258	AX036258 Sequence
38	565.4	88.1	11076	6	AX036260	AX036260 Sequence
39	565.4	88.1	11076	6	AX036262	AX036262 Sequence
40	564.8	88.0	1054	14	HPCPEP63	D00689 Hepatitis C
41	563.8	87.8	1539	6	AR027786	AR027786 Sequence
42	563.8	87.8	1863	6	AR027783	AR027783 Sequence
43	563.8	87.8	3296	14	AB008442	AB008442 Hepatitis
44	563.8	87.8	3296	14	AB008443	AB008443 Hepatitis
45	563.8	87.8	9379	14	HCVPOLYP	AJ000009 Hepatitis

ALIGNMENTS

RESULT 1

A48665
LOCUS A48665 Sequence 3 from Patent WO9604385.
DEFINITION A48665
ACCESSION A48665
VERSION A48665.1 GI:2302378
KEYWORDS
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 642)
AUTHORS Maertens G., Bosman F., De M.G. and Buyse M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 3 15-FEB-1996;

linear PAT 07-MAR-1997

INNOGENETICS NV (BE) Other publication CA 2172273 960215 Other publication AU 3382495 960304. Other publication Location/Qualifiers	AR157324 Sequence 3 from patent US 6245503. AR157324 AR157324 AR157324.1 GI:16218256 Unknown. SOURCE ORGANISM Unclassified.	LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM Unclassified.	642 bp DNA linear PAT 17-OCT-2000
FEATURES source	1..642 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644" 1..639 /note="unnamed protein product" /protein_id="CRA03128.1" /db_xref="GI:2020379" /translation="MPGCSFSLFLALLSCLTIPASAVEVENVSMYVHTDCNSNSI VVEADIMHTGCVPCVRENNSSRWALPTLAARNASVPTTIIRHVDLLVGAAR LCSAMTVGDLCSVFLVSQLFTISPRHETVDDNCISLYPHITGHRMAMDMNWSF TTALVVSQLLRIPQAVDVMVAGHGVLAGLAYYSVWGNWAKVLIVMLLAL"	REFERENCE AUTHORS TITLE JOURNAL FEATURES source	Maertens,G., Bosman,F.; De Martynoff,G. and Buysse,M.-A. Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use Patent: US 6245503-A 3 12-JUN-2001; Location/Qualifiers 1..642 /organism="unknown" BASE COUNT 109 a 195 c 176 g 162 t ORIGIN
mat_peptide	1..636 /product="unnamed"	Query Match 100.0%; Score 642; DB 6; Length 642; Best Local Similarity 100.0%; Pred. No. 1.1e-150; Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
BASE COUNT 109 a 195 c 176 g 162 t ORIGIN			
Query Match 100.0%; Score 642; DB 6; Length 642; Best Local Similarity 100.0%; Pred. No. 1.1e-150; Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy 1 ATGCCCGGTTGCTTTCTCTATCTTCCTCTGGCTTTACTGCTGCTGACCAATTCGA 60 Db 1 ATGCCCGGTTGCTTTCTCTATCTTCCTCTGGCTTTACTGCTGCTGACCAATTCGA 60 Qy 61 GCTTCCCGTTATGAGGTGCGCAACGTTGCCGATGTACCATGTACCAAGACACTGCTCC 120 Db 61 GCTTCCCGTTATGAGGTGCGCAACGTTGCCGATGTACCATGTACCAAGACACTGCTCC 120 Qy 121 AACTCAAGCATTTGTATGAGGACGCGACATCATGCACACCCCGGGTGGGTGCC 180 Db 121 AACTCAAGCATTTGTATGAGGACGCGACATCATGCACACCCCGGGTGGGTGCC 180 Qy 181 TGGCTTCCGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGAGCT 240 Db 181 TGGCTTCCGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGAGCT 240 Qy 241 AGGAACGCGACGCTCCCAACGACATACGACGACGACGACGACGACGACGACGACG 300 Db 241 AGGAACGCGACGCTCCCAACGACATACGACGACGACGACGACGACGACGACGACG 300 Qy 301 GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTGCGGATCTGTCTCTCGTCTCC 360 Db 301 GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTGCGGATCTGTCTCTCGTCTCC 360 Qy 361 CAGCTGTTTACCAATCTCGCTCCCGCATGAGACGCTCAGAGCTGCAATTTGCTCAATC 420 Db 361 CAGCTGTTTACCAATCTCGCTCCCGCATGAGACGCTCAGAGCTGCAATTTGCTCAATC 420 Qy 421 TATCCCGGCCACATAACAGGTCAACGATGCGTGGGATATGATGATGATGATGATGATG 480 Db 421 TATCCCGGCCACATAACAGGTCAACGATGCGTGGGATATGATGATGATGATGATGATG 480 Qy 481 ACAACGGCCCTTGGTATGCGAGCTGCTCGGATCCCAAGCTGCTGGACATGGTG 540 Db 481 ACAACGGCCCTTGGTATGCGAGCTGCTCGGATCCCAAGCTGCTGGACATGGTG 540 Qy 541 GCGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCTTCTCTCTCTCTCTCTCTCTCTCT 600 Db 541 GCGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCTTCTCTCTCTCTCTCTCTCTCTCT 600 Qy 601 GCTAAGGTTTTGATTTGATGCTACTCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCT 642 Db 601 GCTAAGGTTTTGATTTGATGCTACTCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCT 642	Qy 1 ATGCCCGGTTGCTTTCTCTATCTTCCTCTGGCTTTACTGCTGCTGACCAATTCGA 60 Db 1 ATGCCCGGTTGCTTTCTCTATCTTCCTCTGGCTTTACTGCTGCTGACCAATTCGA 60 Qy 61 GCTTCCCGTTATGAGGTGCGCAACGTTGCCGATGTACCATGTACCAAGACACTGCTCC 120 Db 61 GCTTCCCGTTATGAGGTGCGCAACGTTGCCGATGTACCATGTACCAAGACACTGCTCC 120 Qy 121 AACTCAAGCATTTGTATGAGGACGCGACATCATGCACACCCCGGGTGGGTGCC 180 Db 121 AACTCAAGCATTTGTATGAGGACGCGACATCATGCACACCCCGGGTGGGTGCC 180 Qy 181 TGGCTTCCGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGAGCT 240 Db 181 TGGCTTCCGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGAGCT 240 Qy 241 AGGAACGCGACGCTCCCAACGACATACGACGACGACGACGACGACGACGACGACG 300 Db 241 AGGAACGCGACGCTCCCAACGACATACGACGACGACGACGACGACGACGACGACG 300 Qy 301 GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTGCGGATCTGTCTCTCGTCTCC 360 Db 301 GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTGCGGATCTGTCTCTCGTCTCC 360 Qy 361 CAGCTGTTTACCAATCTCGCTCCCGCATGAGACGCTCAGAGCTGCAATTTGCTCAATC 420 Db 361 CAGCTGTTTACCAATCTCGCTCCCGCATGAGACGCTCAGAGCTGCAATTTGCTCAATC 420 Qy 421 TATCCCGGCCACATAACAGGTCAACGATGCGTGGGATATGATGATGATGATGATGATG 480 Db 421 TATCCCGGCCACATAACAGGTCAACGATGCGTGGGATATGATGATGATGATGATGATG 480 Qy 481 ACAACGGCCCTTGGTATGCGAGCTGCTCGGATCCCAAGCTGCTGGACATGGTG 540 Db 481 ACAACGGCCCTTGGTATGCGAGCTGCTCGGATCCCAAGCTGCTGGACATGGTG 540 Qy 541 GCGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCTTCTCTCTCTCTCTCTCTCTCTCT 600 Db 541 GCGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCTTCTCTCTCTCTCTCTCTCTCTCT 600 Qy 601 GCTAAGGTTTTGATTTGATGCTACTCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCT 642 Db 601 GCTAAGGTTTTGATTTGATGCTACTCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCT 642	RESULT 3 LOCUS DEFINITION ACCESSION SEQUENCE 3 FROM PATENT EP1211315. AX452752 AX452752 AX452752 linear DNA 642 bp PAT 06-JUL-2000	

VERSION	AX452752.1										GI:21712437
KEYWORDS	Hepatitis C virus										
SOURCE	Hepatitis C virus										
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.										
REFERENCE	1										
AUTHORS	Maertens,G., Bosman,F., de Martynoff,G. and Buysse,M.A.										
TITLE	Recombinant vectors for producing hcv envelope proteins										
JOURNAL	Patent: EP 1211315-A 3 05-JUN-2002;										
	Innogenetics N.V. (BE)										
FEATURES	Location/Qualifiers										
source	1..642										
	/organism="Hepatitis C virus"										
	/mol_type="genomic DNA"										
	/db_xref="taxon:11103"										
CDS	1..639										
	/note="unnamed protein product"										
	/codon_start=1										
	/protein_id="CAD38066.1"										
	/db_xref="GI:21712438"										
	/translation="MPCGSPSIFLLALLSCLTPASAVEVRNYSVMVHTNDCNSISIVYEADMIHTPGCPVPCVRNNSRCHVALPTPLAARNASVPTTIRRHVDLLVGAARLCSAMYGDLCGVSFLVSPGCTFTISPRRHETVQDCNCSIYGHITGHKRNADMMNWSPTTALVSQLLRIPQVAVDMVAGHWGLAGLYSYMGNWAKVLIIVMLLFL"										
mat_peptide	1..636										
BASE COUNT	109 a 195 c 176 g 162 t										
ORIGIN											
	Query Match 100.0%; Score 642; DB 6; Length 642;										
	Best Local Similarity 100.0%; Pred. No. 1.1e-150;										
	Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
Qy	1	ATGCCCGGTGCTCTTCTCTATCTTCTCTCTTGGCTTTACTGTCTCTGTGACCAATTC	60								
Db	1	ATGCCCGGTGCTCTTCTCTATCTTCTCTTGGCTTTACTGTCTCTGTGACCAATTC	60								
Qy	61	GCTTCCGCTTATGAGGTGCGCAACGTGTCGGGATGACCATGTACAGACGACTGCTCC	120								
Db	61	GCTTCCGCTTATGAGGTGCGCAACGTGTCGGGATGACCATGTACAGACGACTGCTCC	120								
Qy	121	AACTCAGCAATTGTTATGAGCAGCGGACATGATCATGCACACCCCGGTGCGTCC	180								
Db	121	AACTCAGCAATTGTTATGAGCAGCGGACATGATCATGCACACCCCGGTGCGTCC	180								
Qy	181	TGCGTTGGGAGAACAACTCTTCGCGTCTGGGTAGGCTCACCCCGACGTCGACGCT	240								
Db	181	TGCGTTGGGAGAACAACTCTTCGCGTCTGGGTAGGCTCACCCCGACGTCGACGCT	240								
Qy	241	AGGAACGCACGCTCCCAACACACAATACGACGCCACGTCGATTGCTGTTGGGGCG	300								
Db	241	AGGAACGCACGCTCCCAACACACAATACGACGCCACGTCGATTGCTGTTGGGGCG	300								
Qy	301	GCTGCTCTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGCTTCTCGTCTCC	360								
Db	301	GCTGCTCTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGCTTCTCGTCTCC	360								
Qy	361	CAGCTGTTACCATCTCGCTCGCGGACATGAGACGGTGCAGGACTGCAATGTCTCAAT	420								
Db	361	CAGCTGTTACCATCTCGCTCGCGGACATGAGACGGTGCAGGACTGCAATGTCTCAAT	420								
Qy	421	TATCCCGGCACATAACAGGTCACGGTATGCTTGGGATATGATGATGAATGCTGCGCT	480								
Db	421	TATCCCGGCACATAACAGGTCACGGTATGCTTGGGATATGATGATGAATGCTGCGCT	480								
Qy	481	ACAAACGGCCCTTGGTGTATCGACGTCTCCGGATCCCAACAGCTGCTGCGACATGGT	540								
Db	481	ACAAACGGCCCTTGGTGTATCGACGTCTCCGGATCCCAACAGCTGCTGCGACATGGT	540								
Qy	541	CGCGGGGCCCATTTGGGAGTCTCGCGGCTCGCCCTACTATTCCATGGTGGGAAC	600								
Db	541	CGCGGGGCCCATTTGGGAGTCTCGCGGCTCGCCCTACTATTCCATGGTGGGAAC	600								

```

QY 601 GCTAAGGTTTGTGATGTGATGCTACTCTTTGGCTCTCTTAATAG 642
Db 601 GCTAAGGTTTGTGATGTGATGCTACTCTTTGGCTCTCTTAATAG 642

RESULT 4
AX685004
LOCUS AX685004 642 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 3 from Patent WO02055548.
ACCESSION AX685004
VERSION AX685004.1 GI:29371409
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Maertens, G., Bosman, F. and Buyse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 02055548-A 3 18-JUL-2002;
INNOGENETICS N.V. (BE)
FEATURES
source Location/Qualifiers
1..642
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..639
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD06519.1"
/db_xref="GI:29371410"
/translation="MPCGSPSIFLLALLSLCTIPASAYEVRNVSMYHVTVNDCNSGII
VYZAADIMHTPGCVKRENNSSRCWVALPTLAARNASVPVTTIRRHVDLVLGAALA
LCSAMVYGDLCGVSQFLTSPRRHETVQCNCISYPGHITGHRMADWMWMMWSP
TTALVWSOLLRIPOAVVDVMVAGHWGLAGLAYVSMVGNKVLIVMLLPAL"
mat_peptide 1..636
/product="unnamed"
BASE COUNT 109 a 195 c 176 g 162 t
ORIGIN

Query Match 100.0%; Score 642; DB 6; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.1e-150;
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCCGGTTGCTCTTTCTCTATCTTCTCTGTGGCTTTACTGTCTCTGACCAATTCGA 60
Db 1 ATGCCCGGTTGCTCTTTCTCTATCTTCTCTGTGGCTTTACTGTCTCTGACCAATTCGA 60

QY 61 GCTTCCGGTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTACGACAGCTGTCTCC 120
Db 61 GCTTCCGGTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTGTACGACAGCTGTCTCC 120

QY 121 AACTCAAGCATTTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTGCGCC 180
Db 121 AACTCAAGCATTTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTGCGCC 180

QY 181 TGC GTTTCGGGAGAACAACTCTTCCCGTGTGGGTAGCGGTTCACCCCGACGGTCGCAGCT 240
Db 181 TGC GTTTCGGGAGAACAACTCTTCCCGTGTGGGTAGCGGTTCACCCCGACGGTCGCAGCT 240

QY 241 AGGAACGCGACGTCGCCACCAACGACAAATACGACGCCACGTCGATTTGCTGTTGGGGCG 300
Db 241 AGGAACGCGACGTCGCCACCAACGACAAATACGACGCCACGTCGATTTGCTGTTGGGGCG 300

QY 301 GCTGCTCTGTGTTCCGGTGTATGATGCTGGGGATCTCTGCGGATCTGTCCTCGTCTCC 360
Db 301 GCTGCTCTGTGTTCCGGTGTATGATGCTGGGGATCTCTGCGGATCTGTCCTCGTCTCC 360

QY 361 CAGCTGTTTACCATTCTCCGCTTCGCCGCGCATGAGACGGTGCAGGACTGCAATTTGCTCAATC 420
Db 361 CAGCTGTTTACCATTCTCCGCTTCGCCGCGCATGAGACGGTGCAGGACTGCAATTTGCTCAATC 420

```


b	335	GCCTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTA	394			
y	242	GGAACGCCAGCGTCCCACACAGACAATACGACGCCACGTCGATTGCTCGTTGGGGCGG	301			
b	395	GGAACGCCAGCGTCCCACACAGACAATACGACGCCACGTCGATTGCTCGTTGGGGCGG	454			
y	302	CTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTTGCGGATCTGTCTCTCGTCTCC	361			
b	455	CTGCTTCTGTTCCGCTATGTACGTGGGGGACCTCTGCGGATCTGTCTCTCGTCTCC	514			
y	362	AGCTGTTACCATCTCGCTCGCGGCATGAGACGCTGACGAGACTGCAATTGCTCAATCT	421			
b	515	AGCTGTTACCATCTCGCTCGCGGCATGAGACGCTGACGAGACTGCAATTGCTCAATCT	574			
y	422	ATCCCGGCCACATAAAGGTCACCGTATGCTTGGGATATGATGATCACTGGTCGCCTA	481			
b	575	ATCCCGGCCACATAAAGGTCACCGTATGCTTGGGATATGATGATCACTGGTCGCCTA	634			
y	482	CACGCGCTCGTGGTATCCGAGCTGCTCCGGATCCCCACAAGCTGTGTCGACATGTGG	541			
b	635	CACGCGCTCGTGGTATCCGAGCTGCTCCGGATCCCCACAAGCTGTGTCGACATGTGG	694			
y	542	CGGGGGCCATTTGGGGAGTCTCTGCGGGCTCGGCTACTATTCATGTCGGGAACCTGGG	601			
b	695	CGGGGGCCATTTGGGGAGTCTCTGCGGGCTCGGCTACTATTCATGTCGGGAACCTGGG	754			
y	602	CTAAGTTTTCATGCTGATGCTACTCTTTGCTCTCTAATAG	642			
b	755	CTAAGTTTTCATGCTGATGCTACTCTTTGCTCTCTAATAG	795			
<hr/>						
RESULT 6						
AR157325						
LOCUS	AR157325	795 bp	DNA linear PAT 17-OCT-2001			
DEFINITION	Sequence 5 from patent US 6245503.					
ACCESSION	AR157325					
VERSION	AR157325.1	GI:16218258				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 795)					
AUTHORS	Maericens,G., Bosman,F., De Martynoff,G. and Buysee,M.-A.					
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use					
JOURNAL	Patent: US 6245503-A 5 12-JUN-2001;					
FEATURES	Location/Qualifiers 1..795 /organism="unknown"					
BASE COUNT	130 a	240 c	231 g 194 t			
ORIGIN						
<hr/>						
Query Match 97.9%; Score 628.2; DB 6; Length 795;						
Best Local Similarity 98.8%; Pred.No. 3.2e-147;						
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;						
QY	2	TGCCCGGTTGCTCTTCTCTATCTTCCTCTTGCGTTTACTGTCTGTCTGACCATTCCAG	61			
Db	155	TGCCCGGTTGCTCTTCTCTATCTTCCTCTTGCGTTTACTGTCTGTCTGACCGTTCCAG	214			
QY	62	CTTCGGCTTATGAGTGGGAACTGTCGGGATGTACCATGTCCAGAACGACTGCTCCA	121			
Db	215	CTTCGGCTTATGAGTGGGAACTGTCGGGATGTACCATGTCCAGAACGACTGCTCCA	274			
QY	122	ACTCAAGCATTTGTATGAGCAGCGGATGATCATGCACCCCCGGGTGCGTGCCT	181			
Db	215	ACTCAAGCATTTGTATGAGCAGCGGATGATCATGCACCCCCGGGTGCGTGCCT	334			
QY	182	GCGTTCCGGGAGAACAACTCTTCCGCTGCTGGTAGCGCTCACCCCCCAACGCTTCGAGCTA	241			
Db	335	GCGTTCCGGGAGAACAACTCTTCCGCTGCTGGTAGCGCTCACCCCCCAACGCTTCGAGCTA	394			

QY	421	TATCCCGGCCACATAACAGGTCACCGTATGGCTTGGGATATGATGAAGTGGTGGCCT	480		
Db	421	TATCCCGGCCACATAACAGGTCACCGTATGGCTTGGGATATGATGAAGTGGTGGCCT	480		
QY	481	ACAAAGCGCCCTGGTGGTATCCAGCTGCTCCGGATCCCAAGCTGTCGTGGACATGGT	540		
Db	481	ACAAAGCGCCCTGGTGGTATCCAGCTGCTCCGGATCCCAAGCTGTCGTGGACATGGT	540		
QY	541	GCAGGGGGCCCATTTGGGGAGTCTCGGGGCTCGCTACTACTTCCATGCTGGGGAAGTGG	600		
Db	541	GCAGGGGGCCCATTTGGGGAGTCTCGGGGCTCGCTACTACTTCCATGCTGGGGAAGTGG	600		
QY	601	GCTAAGGTTTGGATGTGATGCTACTCTTTGCTCTCTCTAATAG	642		
Db	601	GCTAAGGTTTGGATGTGATGCTACTCTTTGCTCTCTCTAATAG	642		
RESULT 5					
A48667	A48667	795 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	Sequence 5 from Patent WO9604385.				
DEFINITION	A48667				
ACCESSION	A48667.1				
VERSION	GI:2302380				
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 795)				
AUTHORS	Maertens, G., Bosman, F., De, M. G. and Buyse, M.				
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE				
JOURNAL	Patent: WO 9604385-A 5 15-FEB-1996;				
COMMENT	INNOGENETICS NV (BE)				
FEATURES	Other publication CA 2172273 960215				
	Other publication AU 3382495 960304.				
source	Location/Qualifiers				
	1..795				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
CDS	1..792				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAA03129.1"				
	/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLVLDG				
mat_peptide	VNYATGNLPGCSFSLILALISCLTPASAYEVNVMGYHTVDLNSGSIYVYADAM				
	IMTPTGCVPCVRENNSSRCWALPTLAARNASVPTTIRHVDLLVGAFAAFCSAMTV				
	GDLGCVFLVSLQFLTISPRHETVQDCNSIYPGHITGHRMAMDMNNWSPPTLALVVS				
	QLLRIPQAVDVMVAGAHGVLAGLAYISVMGNWAKVLIVMLLFP"				
	1..789				
BASE COUNT	130 a	240 c	231 g	194 t	
ORIGIN					
Query Match	97.9%; Score 628.2; DB 6; Length 795;				
Best Local Similarity	98.8%; Pred. No. 3.2e-147;				
Matches	633; Conservative	0; Mismatches	8; Indels	0; Gaps	0;
QY	2	TGCCCCGGTTGGCTCTTTCTCTATCTTCTCTCTTGGCTTTACTGTCCTGTGTGACCATTCAC	61		
Db	155	TGCCCCGGTTGGCTCTTTCTCTATCTTCTCTTGGCTTTGTGTGTGTGTGACCGTTCCAG	214		
QY	62	CTTCCCGCTTATGAGTGGCGCAAGCTGTCGGGATGATACCATGTGCACGAACGACTGCTCCA	121		
Db	215	CTTCCCGCTTATGAGTGGCGCAAGCTGTCGGGATGATACCATGTGCACGAACGACTGCTCCA	274		
QY	122	ACTCAGACATTTGTTATGAGCAGCGGACATGATCATGCACACCCCGGGTGGTGGCCT	181		
Db	275	ACTCAGACATTTGTTATGAGCAGCGGACATGATCATGCACACCCCGGGTGGTGGCCT	334		
QY	182	CGCTTCGGGAGAACAACTCTTCGGCTGTGGGTAGCGGTCCACCCCAACGCTCGCAGCTA	241		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 177.486 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATGCCCGTGTCTTCTC.....TACTTTGCTCTCTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642	100.0	642	17	AAT12704
2	642	100.0	642	24	AA148913
3	628.2	97.9	795	17	AAT12705
4	628.2	97.9	795	24	AA148914
5	619.8	96.5	2082	24	AA148919
6	619.8	96.5	2086	17	AAT12973
7	619.8	96.5	2433	17	AAT12974
8	608.8	94.8	2434	24	AA148940

9	566.4	88.2	3461	15	AAQ64068	Non-A, non-B hepat
10	566.4	88.2	3461	16	AAT30386	5'UTR/CORE/ENV/NSI
11	565.4	88.1	1880	13	AAQ24467	NANB hepatitis vir
12	565.4	88.1	2187	19	ABA03491	Cuticle protein 1
13	565.4	88.1	2540	14	AAQ43889	NANB hepatitis vir
14	565.4	88.1	2540	15	AAQ63753	NANBHV genomic fra
15	565.4	88.1	9605	24	ABK91411	Hepatitis C virus
16	565.4	88.1	9605	24	ABK91424	Hepatitis C virus
17	565.4	88.1	9605	24	ABK91425	Hepatitis C virus
18	565.4	88.1	9605	24	ABK91426	Hepatitis C virus
19	565.4	88.1	9605	24	ABK91428	Hepatitis C virus
20	565.4	88.1	9605	24	ABK91429	Hepatitis C virus
21	565.4	88.1	9605	24	ABK91430	Hepatitis C virus
22	565.4	88.1	9605	24	ABK91431	Hepatitis C virus
23	565.4	88.1	9605	24	ABK91432	Hepatitis C virus
24	565.4	88.1	9605	24	ABK91433	Hepatitis C virus
25	565.4	88.1	9605	24	ABK91434	Hepatitis C virus
26	565.4	88.1	9608	24	ABK91427	Hepatitis C virus
27	565.4	88.1	11062	24	AAQ25331	Hepatitis C virus
28	565.4	88.1	11076	21	AAA98965	Hepatitis C virus
29	563.2	87.7	1251	13	AAQ15363	Fragment of NANB h
30	563.2	87.7	1251	13	AAQ26981	HCV gene 1. Hepat
31	563.2	87.7	3360	17	AAT03677	Hepatitis C genome
32	563.2	87.7	9413	16	AAT03960	Partial HCV non-sc
33	563.2	87.7	9413	16	AAQ81559	Hepatitis C virus
34	563.2	87.7	9413	16	AAQ80498	DNA encoding HCV p
35	563.2	87.7	9413	24	AAQ25517	Hepatitis C virus
36	563.2	87.7	9413	25	AAQ49655	Hepatitis C virus
37	563.2	87.7	9413	25	AAQ53723	Hepatitis C virus
38	562.2	87.6	2540	13	AAQ29628	Hepatitis C virus
39	560	87.2	1251	13	AAQ25610	HCV in expression
40	560	87.2	1398	13	AAQ22140	Hepatitis C virus
41	559	87.1	742	13	AAQ20926	C10-E15 DNA fragme
42	559	87.1	932	13	AAQ20923	C10-E12 DNA fragme
43	559	87.1	1562	19	AAV60672	Fragment #5 isolat
44	559	87.1	1953	25	AAQ55222	Plasmid pIDK2 DNA
45	559	87.1	2829	19	AAV60673	Fragment #6 isolat

ALIGNMENTS

RESULT 1

AAT12704

ID AAT12704 standard; DNA; 642 BP.

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

XX AC AAT12704;

proteins - in presence of di:sulphide bond cleavage agent, to
produce proteins suitable for direct use in vaccines or diagnostic
assays of HCV

Claim 23; Fig 21; 146pp; English.

AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
and E2 protein coding sequence constructs. These sequences are included
in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
The recombinant proteins can then be isolated using a method of the
invention. In the method, the envelope proteins are purified by
carrying out a disulphide bond cleavage, or a reduction step with a
disulphide bond cleavage agent, after lysis of recombinant host cells.
The constructs containing the purified HCV envelope proteins can be used
for vaccinating humans against HCV, for in vitro detection of HCV
antibodies in a sample, and in a serotyping assay for detecting one or
more serological types of HCV present in a biological sample. The
constructs can also be immobilised on a solid substrate and incorporated
into a reversed phase hybridisation assay for determining the presence or
the genotype of HCV. The new purification method preserves the
conformation of the recombinantly expressed E1, E2 and E1/E2, and
eliminates contaminating proteins. Antigens isolated using this method
are more reactive with human sera than those isolated by known
techniques.

Sequence 642 BP; 109 A; 195 C; 176 G; 162 T; 0 other;

Query Match	100.0%	Score 642	DB 17	Length 642
Best Local Similarity	100.0%	Mismatches 0	Indels 0	Gaps 0
Matches 642	Conservative			
QY	1	ATGCCCGGTGCTCTTTCTCTATCTCTCTTTGGCTTTACTGTCTGTCTGACCATTTCCA	60	
DB	1	ATGCCCGGTGCTCTTTCTCTATCTCTCTTTGGCTTTACTGTCTGTCTGACCATTTCCA	60	
QY	61	GCTTCCGCTTATGAGGTGCGCAACGTGTCCGGATGTACCATGTCAAGAACACTGCTCC	120	
DB	61	GCTTCCGCTTATGAGGTGCGCAACGTGTCCGGATGTACCATGTCAAGAACACTGCTCC	120	
QY	121	AACTCAAGCATTTGTATGAGGCACGCGACATGATCTGCACACCCCGGTGTGGTGCCT	180	
DB	121	AACTCAAGCATTTGTATGAGGCACGCGACATGATCTGCACACCCCGGTGTGGTGCCT	180	
QY	181	TGCGTTCCGGAGAACAACTCTTCCTCGCTGCTCGGTAGCGCTCACCCCAACGCTCGCAGCT	240	
DB	181	TGCGTTCCGGAGAACAACTCTTCCTCGCTGCTCGGTAGCGCTCACCCCAACGCTCGCAGCT	240	
QY	241	AGGAACGCCAGCGGTCCCCACCAAGCAATPACGACGCCAGCTCGATTTGCTTGTGGGGCG	300	
DB	241	AGGAACGCCAGCGTCCCCACCAAGCAATPACGACGCCAGCTCGATTTGCTTGTGGGGCG	300	
QY	301	GCTGCTCTCTGTTTCGCTATGTACGTGGGGGATCTCTGCGATCTGTCTTCCTCGTCTCC	360	
DB	301	GCTGCTCTCTGTTTCGCTATGTACGTGGGGGATCTCTGCGATCTGTCTTCCTCGTCTCC	360	
QY	361	CAGCTGTTTACCACTCTCGCTCCCGGCATGAGACCGGTGAGACTGCAATTCCTCAATC	420	
DB	361	CAGCTGTTTACCACTCTCGCTCCCGGCATGAGACCGGTGAGACTGCAATTCCTCAATC	420	
QY	421	TATTCGGGCACATACAGGTACCGGTATGGCTTGGGATATGATGATGAATCGTGCCT	480	
DB	421	TATTCGGGCACATACAGGTACCGGTATGGCTTGGGATATGATGATGAATCGTGCCT	480	
QY	481	ACAAACGCCCTTGGTGTATCGAGCTGCTCGGATCCCAAGCTGTCGTGGACATGGTG	540	
DB	481	ACAAACGCCCTTGGTGTATCGAGCTGCTCGGATCCCAAGCTGTCGTGGACATGGTG	540	
QY	541	GCGGGGGCCATTTGGGAGTCTCGCGGGCTCGCTACTATTTCATGTGGGGAACTGG	600	
DB	541	GCGGGGGCCATTTGGGAGTCTCGCGGGCTCGCTACTATTTCATGTGGGGAACTGG	600	
QY	601	GCTAAGGTTTTGGATTGTGATGCTACTCTTTGCTCTCTAATAG	642	

501 CCTAAGGTTTGAATGTGATGCTACTCTTTGCTCTCTAATAG 642

RESULT 2	
AAL48913	AAL48913 standard; DNA; 642 BP.
ID	
XX	
XX	AAL48913;
XX	
XX	24-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus clone HCCI9A E1 protein coding sequence.
XX	
XX	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW	viricide; immunostimulant; vaccine; ds.
KX	
XX	Hepatitis C virus.
OS	
XX	
FN	WO200255548-A2.
XX	
FD	18-JUL-2002.
XX	
PF	11-JAN-2002; 2002WO-EPO0219.
XX	
PR	11-JAN-2001; 2001US-260699P.
PR	30-AUG-2001; 2001US-315768P.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Maertens G, Bosman F, Buysse M;
XX	
DR	WPI; 2002-599657/64.
DR	P-FSDB; AAO18660.
XX	
XX	New therapeutic vaccine compositions comprising at least one purified recombinant hepatitis C virus (HCV) single or specific oligomer; PT recombination envelope protein E1 or E2, useful for immunizing hum PT from HCV infection -
XX	
PS	Example 2; Page 158-159; 243pp; English.
XX	
CC	The present invention relates to new therapeutic vaccine compos CC inducing hepatitis C virus (HCV)-specific antibodies, compris CC composition containing at least one purified recombinant HCV si CC specific oligomeric recombinant envelope proteins selected from CC an E2 protein, and optionally a pharmaceutical adjuvant. The va CC useful for inducing HCV-specific antibodies or for immunizing h CC against HCV. The recombinant HCV E1 and/or E2 proteins are usef CC vaccines or therapeutics, in HCV screening and confirmatory ant CC tests, for raising antibodies, in the preparation of medicame CC in vitro monitoring of HCV disease or prognosing the response t CC treatment of patients suffering from HCV infection. The present CC is a coding sequence described in the exemplification of the in
XX	
SQ	Sequence 642 BP; 109 A; 195 C; 176 G; 162 T; 0 other;
	Query Match 100.0%; Score 642; DB 24; Length 642;
	Best Local Similarity 100.0%; Pred. No. 1.3e-174;
	Matches 642; Conservative 0; Mismatches 0; Indels 0;
Qy	1 ATGCCCGGTGCTTTTCTATCTTCCTTGCTTTACGTCTGTCTGTCTGCAC
Db	1 ATGCCCGGTGCTTTTCTATCTTCCTTGCTTTACGTCTGTCTGTCTGCAC
Qy	61 GCTTCCGCTTAGAGTGGCGAACGTGTCGGGATGTACCATGTCACGAACA
Db	61 GCTTCCGCTTAGAGTGGCGAACGTGTCGGGATGTACCATGTCACGAACA
Qy	121 AACCTCAAGCATTTGTATCAGGCAGCGGACATGATGCACACCCCGGTG
Db	121 AACCTCAAGCATTTGTATCAGGCAGCGGACATGATGCACACCCCGGTG
Qy	181 TGCGTTCCGGAGAACAACTCTTCCCCTGCTGGGTAGCGCTCACCCCCACGCT


```
Db 181 TGGTTCGGGAGAACAACTCTTCCGCTGCTGGGTAGCGTCAACCCACGCTCGCAGCT 240
Qy 241 AGGAACGCCAGCGTCCACACGACAAATACGACGCGCACGTCGATTTGCTCGTTGGGGCG 300
Db 241 AGGAACGCCAGCGTCCACACGACAAATACGACGCGCACGTCGATTTGCTCGTTGGGGCG 300
Qy 301 GCTGCTCTCTGTTCCGCTATGATGAGTGGGGATCTCGGGATCTGCTTCCTCGTCTCC 360
Db 301 GCTGCTCTCTGTTCCGCTATGATGAGTGGGGATCTCTCGGGATCTGCTTCCTCGTCTCC 360
Qy 361 CAGCTGTTACCACTCTCGCTCGCGGCATGAGACGCTGAGGAGTCAATTTGCTCAATC 420
Db 361 CAGCTGTTACCACTCTCGCTCGCGGCATGAGACGCTGAGGAGTCAATTTGCTCAATC 420
Qy 421 TATCCCGGCCACATAACAGGTCAACGCTATGCTGGGATATGATGATGAACCTGCGCCT 480
Db 421 TATCCCGGCCACATAACAGGTCAACGCTATGCTGGGATATGATGATGAACCTGCGCCT 480
Qy 481 ACAACGCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 ACAACGCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy 541 GCGGGGGCCATTGGGGAGTCTCGCGGGCTCGCCTACTATTCCATGTTGGGGAAGTGG 600
Db 541 GCGGGGGCCATTGGGGAGTCTCGCGGGCTCGCCTACTATTCCATGTTGGGGAAGTGG 600
Qy 601 GCTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 642
Db 601 GCTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 642
```

RESULT 3

```
AAT12705
ID AAT12705 standard; DNA; 795 BP.
XX
AC AAT12705;
XX
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC110A.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of disulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2-protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
```

```
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
```

SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match 97.9%; Score 628.2; DB 17; Length 795;

Best Local Similarity 98.8%; Pred. No. 1.3e-170; Length 795;

Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy 2 TGGCCGGTTGCTCTTTCTCTATCTTCTTGGCTTTTACTGTCTGTCTGACCAATTCAG 61
Db 155 TGGCCGGTTGCTCTTTCTCTATCTTCTTGGCTTTTACTGTCTGTCTGACCGTTCAG 214
Qy 62 CTTCGGCTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTCCAGAACGACTGTCCA 121
Db 215 CTTCGGCTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTCCAGAACGACTGTCCA 274
Qy 122 ACTCAAGCATTTGTATGAGGACGCGACATGATGCACACCCCGGGTGGCGCCT 181
Db 275 ACTCAAGCATTTGTATGAGGACGCGACATGATGCACACCCCGGGTGGCGCCT 334
Qy 182 GCGTTCCGGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCACGCTCCAGCTA 241
Db 335 GCGTTCCGGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCACGCTCCAGCTA 394
Qy 242 GGAACGCCAGCGTCCACCAACGACAAATACGACGCGACGTCGATTTGCTGGGGCGG 301
Db 395 GGAACGCCAGCGTCCACCAACGACAAATACGACGCGACGTCGATTTGCTGGGGCGG 454
Qy 302 CTGCTCTCTGCTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 361
Db 455 CTGCTCTCTGCTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 514
Qy 362 AGCTGTTTCAACCTCTCGCGCATGAGACGCTGAGACGCTGAGACGCTGAGACGCTGAG 421
Db 515 AGCTGTTTCAACCTCTCGCGCATGAGACGCTGAGACGCTGAGACGCTGAGACGCTGAG 574
Qy 422 ATCCCGGCCACATAACAGGTCAACGATGATGATGATGATGATGATGATGATGATGATG 481
Db 575 ATCCCGGCCACATAACAGGTCAACGATGATGATGATGATGATGATGATGATGATGATG 634
Qy 482 CAACGCCCTGCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
Db 635 CAACGCCCTGCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 694
Qy 542 CGGGGGCCCATTTGGGAGTCTCGCGGGCTCGCCTACTATTCCATGTTGGGGAAGTGG 601
Db 695 CGGGGGCCCATTTGGGAGTCTCGCGGGCTCGCCTACTATTCCATGTTGGGGAAGTGG 754
Qy 602 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 642
Db 755 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 795
```

RESULT 4

```
AAL48914
ID AAL48914 standard; DNA; 795 BP.
XX
AC AAL48914;
XX
DT 24-OCT-2002 (first entry)
```


Db	932	ACTCAAGATTATTTGGTATGAGGCAGCGACGTGATCATGCACACCCCGGGTGGTGCCT	991
Qy	182	GC GTTCCGGAGAA CAAC TCTTCCCGCTGCTGGGTAGGCTCACCCACGCTCGCAGCTA	241
Db	992	GC GTTCCGGAGAGACAA TTTCTCCCGCTGCTGGGTAGGCTCACTCCACGCTCGCGCCA	1051
Qy	242	GGAA CGCCAGCGTCCCCACACGACACAATACGACGCCACGTCGATTTGCTGTTGGGGCGG	301
Db	1052	GA AACAGCAGCATCCCCACATCAAGACAATACGACGCCATGTCGATTTGCTGTTGGGGCAG	1111
Qy	302	CTGCTCTCTGTTCCGCTPATGATCGTGGGGATCTTCGGGATCTGCTCTCTGCTCTCCC	361
Db	1112	CTGCTCTCTGCTCGCCATGATGCTGGGGGATCTCTCGGATCTGTCTTCTCTGCTCTCCC	1171
Qy	362	AGCTGTTACCATCTCGCTCGCCGATCAGACGGTGCAGGACTGCAATGCTCAATCT	421
Db	1172	AGCTGTTACCATCTCGCTCGCCGATCAGACGGTGCAGGACTGCAATGCTCAATCT	1231
Qy	422	ATCCCGGCCACATAACAGGTCAACGTTATGGCTTGGGATATGATGATGAAC TGTGCGCTA	481
Db	1232	ATCCCGGCCACGTTCAAGTCAACCGATGGCTTGGGATATGATGATGAAC TGTGCGCTA	1291
Qy	482	CAACGGGCCCTGGTGGTATCGCAGCTGCTCGGATCCACACAGCTGCTGTCGACATGTGG	541
Db	1292	CAACAGCCCTGGTGGTATCGCAGTTACTCGGATCCCAACAGCCATCGTGCACATGTGG	1351
Qy	542	CGGGGGCCATTTGGGGAGTCCCTGGCGGGCTTCGCTACTATTCATCGTGGGGAAC TGGG	601
Db	1352	CAGGGGCCCATTTGGGGAGTCCCTGGCGGGCTTCGCTACTATTCATCGTGGGGAAC TGGG	1411
Qy	602	CTAAGGTTTCATTTGTGATGCTACTCTTTGCT	633
Db	1412	CTAAGGTTTCATTTGTGATGCTACTCTTTGCT	1443

RESULT 10
AAT30386
ID AAT30386 standard; cDNA; 3461 BP.
XX
XX AAT30386;
AC
XX
XX 22-AUG-1996 (first entry)
DT
DT
XX
DE 5'UTR/CORE/ENV/NS1/NS2/NS3 cDNA from HCV (#4).
XX
XX Hepatitis C virus; HCV; antigen; detection; antibody; ds.
XX
XX Hepatitis C virus.
OS

XX	Key	Location/Qualifiers
PH	CDS	307..3461
FT		/*tag= a
FT		307..879
FT	misc_feature	/*tag= b
FT		/product= Core peptide
FT		880..1455
FT	misc_feature	/*tag= c
FT		/product= ENV1
FT		1456..2736
FT	misc_feature	/*tag= d
FT		/product= NS1/ENV2
FT		2737..3461
FT	misc_feature	/*tag= e
FT		/product= NS2 and NS3

XX PN JP07133291-A.

XX
PD 23-MAY-1995.

XX
PE 18-JUN-1993: 93JP-0147944.

XX
10 JUN 1983
93JP-0147944

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

XX	(TOFU) TONEN CORP.
PA	
XX	WPI; 1995-220780/29.
XX	P-PSDB; AAR98361.
XX	
PT	Recombinant polypeptide comprising partial NS1 region of hepatitis
PT	non-A non-B viral antigen - used in a method for detecting
PT	antibodies against hepatitis non-A non-B virus.
XX	
XX	Disclosure; Page 10-12; 15pp; Japanese.
PS	
XX	
CC	The sequences given in AAT30386-87 encode the 5'UTR/CORE/ENV/NS1/NS2/
CC	NS3 protein region derived from hepatitis C virus (HCV) isolates #4
CC	and #6 respectively. The proteins encoded by these sequences partic.
CC	encode amino acids 384-495 of the HCV NS1 antigen. These protein
CC	fragments may be used in the detection of antibodies against HCV.
CC	
XX	
XX	Sequence 3461 BP; 638 A; 1046 C; 1012 G; 765 T; 0 other;
XX	
XX	

Query Match	88.2%	Score 566.4	DB 16	Length 3461
Best Local Similarity	93.5%	Pred. No. 1.4e-152		
Matches 591	Conservative 0	Mismatches 41	Indels 0	Gaps 0
QY	2	TGCGCGTTCCTTTCTCTATCTTCCTCTGCGCTTTACTGTCTCTGTCGACCATTCACG 61		
Db	812			
QY	62	CTTCCGCTTATGAGGTGCGCAACGCTGTCGGGATATACCATGTACGACGACGCTGCTCA 121		
Db	872	CTTCCGCTTATGAATGCGCAACGCTGTCGGGGGTGACCATGTACAAAGACTGCTCCA 931		
QY	122	ACTCAAGCATTTGCTATGAGGACGGGACATGATCATGCACACCCCCGGGTGCGTCCCT 181		
Db	932	ACTCAAGTATTTGCTATGAGGACGGGACGTGATCATGCACACCCCCGGGTGCGTCCCT 991		
QY	182	CGCTTCGGGAGAACCACTCTTCCCGTGTCTGGGTAGCGTCAACCCGACGCTCGCAGCTA 241		
Db	992	CGCTTCGGGAGACAAATTTCTCCCGTGTCTGGGTAGCGTCACTCCGACGCTCGCGGCCA 1051		
QY	242	GGAAGCCAGCGTCCCCACACGACAAATACGACGCCAGCTGCAATTTGCTCGTTGGGGCGG 301		
Db	1052	GAACACAGCAGCATCCCCACATACGACATACGACGCCATGTCGATTTGCTGTTGGGGCAG 1111		
QY	302	CTGCTCTCTGTTCCGCTATGTACGTGGGGATCTTGCGGATCTGTCTTCTCTGCTCTCCC 361		
Db	1112	CTGCTCTCTGCTCCGCCATGTACGTGGGGATCTCTCGGATCTGTCTTCTCTGCTCTCCC 1171		
QY	362	AGCTGTTACCATCTCGCTCGCGGCATACGACGGTGCAGGACTGCAAATGCTCAATCT 421		
Db	1172	AGCTGTTACCTTTCTACCTCGCGGTATGACAGCGTACAGGACTGCAACTGCTCAATCT 1231		
QY	422	ATCCCGGCCACATAACAGGTCAACGGTATGCGTTCGGGATATGATGAACTGTCGCGCTA 481		
Db	1232	ATCCCGGCCACGTGTCAAGTCAACCGCATGGCTTGGGATATGATGAACTGTGCGCTA 1291		
QY	482	CAACGCCCTTGGTGTATCGAGCTGCTCCGGATCCGACAGCTGTCGTGGACATGTGTGG 541		
Db	1292	CAACAGCCCTGGTGTATCGAGTTACTCCGGATCCCAAGCCATCGTGGACATGTGTGG 1351		
QY	542	CGGGGGCCCATGGGGAGTCTCTGGCGGCGCTCGCCTACTATTCATGGTGGGGAACCTGGG 601		
Db	1352	CAGGGGCCCATGGGGAGTCTCTGGCGGCGCTTGCTACTATTCATGGTGGGGAACCTGGG 1411		
QY	602	CTAAGGTTTGGATGTGATGCTACTCTTTGCT 633		
Db	1412	CTAAGGTCTTGATGTGATGCTACTCTTTGCT 1443		

RESULT 11
AAQ24467
ID AAQ24467 standard; DNA; 1880 BP.
XX

Mon Dec 22 13:28:48 2003

us-09-899-303a-3.rng

Best Local Similarity 93.5%; Pred. No. 2.2e-152;
Matches 590; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 TGCCCGGTGGCTTTCTCTATCTTCTCTTGGCTTTACTGTCCTGTCTGACCATTCAG 61
DB 506 TGCCCGGTGGCTTTCTCTATCTTCTCTTGGCTTTGTCCTGTGTGACCATTCAG 565

QY 62 CTTCCCGTTATAGGTGCGCAAGCTGTCCGGGATGTACCATGTTCAGAACTGCTCCA 121
DB 566 CTTCCCGTTATAGGTGCGCAAGCTGTCCGGGATATACCATGTTCAGAACTGCTCCA 625

QY 122 ACTCAAGCATTTGTATGAGCGAGCGGACATGATGACACACCCCGGTGGCTTCC 181
DB 626 ACTCAAGCATTTGTATGAGCGAGCGGACATGATGACACACCCCGGTGGCTTCC 685

QY 182 CGTTTCGGGAGAACATCTTCCCGTCTGGGTAGCGGTTCACCCCGCTCGCAGTCA 241
DB 686 CGTTTCGGGAGAACATCTTCCCGTCTGGGTAGCGGTTCACCCCGCTCGCAGTCA 745

QY 242 GGAACGCCAGGTCCCGACACGACAAATACGACGCCACGTGCGATTGCTGTTGGGCGG 301
DB 746 GGAATGCCAGGTCCCGACACGACAAATACGACGCCACGTGCGATTGCTGTTGGGCGG 805

QY 302 CTGCTCTCTGCTTCGCTATGTACGTGGGGATCTCTCGGATCTGTTCTCTGCTCCC 361
DB 806 CTGCTCTCTGCTTCGCTATGTACGTGGGGATCTCTCGGATCTGTTCTCTGCTCCC 865

QY 362 AGCTGTTCCACCATCTCGGCTCGCGGATGAGCGGTGCAGGACTGCAATGTCTCAATCT 421
DB 866 AGCTGTTCCACCATCTCGGCTCGCGGATGAGCAGTGCAGGACTGCAATGTCTCAATCT 925

QY 422 ATCCCGGCCACATAACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 481
DB 926 ATCCCGGCCATTTATCAGGTACCGCATGCTTGGGATATGATGATGATGATGATGATGAT 985

QY 482 CAAAGCGCTTGTGTATGCGAGTCTCGGATCCACAGCTGTGCGTGGACATGTTGG 541
DB 986 CAAAGCGCTTGTGTATGCGAGTCTCGGATCCACAGCTGTGCGTGGACATGTTGG 1045

QY 542 CGGGGGCCCATTTGGGGAGTCTTGGGGGCTTGGGGCTTGGCTTCTTCCATTTGGGAACTGGG 601
DB 1046 CGGGGGCCCATTTGGGGAGTCTTGGGGGCTTGGGGCTTGGCTTCTTCCATTTGGGAACTGGG 1105

QY 602 CTAAGGTTTGTATGTGCTACTCTTTGC 632
DB 1106 CTAAGGTTCTGATTGTGGGCTACTCTTGC 1136

RESULT 13
AAQ43889 standard; cDNA to mRNA; 2540 BP.
ID AAQ43889
XX AC AAQ43889;
XX DT 21-OCT-1993 (first entry)
XX DE NANB hepatitis virus polynucleotide N-2540-2.
XX KW Non-A, non-B; virus; polymerase chain reaction; detection;
KW sensitive; specific; HCV; NANBH; ss.
XX OS Non-A, non-B hepatitis virus.
XX Key Location/Qualifiers
FH CDS 342..2540
FT /*tag= a
FT 1..341
FT /*tag= b
FT /note= "from 5' terminal of NANBH virus RNA"
XX JP05091884-A.
XX 16-APR-1993.
PD

XX 10-APR-1991; 91JP-0196175.
XX 12-JUN-1990; 90JP-0153401.
XX 08-NOV-1990; 90JP-0304405.
XX (NAKA/) NAKAMURA T.
XX WPI; 1993-199637/25.
XX P-PSDB; AAR38279.
XX Antigen related to non-A and non-B hepatitis virus - comprises
PT non-translation region comprising 340 - 341 moles. of nucleotides,
PT non-translation region comprising 1885 - 2551 moles. of
PT nucleotides including region 1,149 and, etc.
XX Claim 3; Page 19-20; 73pp; Japanese.
XX The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
CC which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
CC The polypeptide it encodes may be used in a system for detecting
CC NANB hepatitis. This method is highly specific and sensitive, and
CC can detect NANB hepatitis virus which could not be detected by
CC conventional methods.
XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;
SQ

Query Match 88.1%; Score 565.4; DB 14; Length 2540;
Best Local Similarity 93.5%; Pred. No. 2.4e-152;
Matches 590; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 TGCCCGGTGGCTTTCTCTATCTTCTCTTGGCTTTACTGTCCTGTCTGACCATTCAG 61
DB 847 TGCCCGGTGGCTTTCTCTATCTTCTCTTGGCTTTGTCCTGTCTGACCATTCAG 906

QY 62 CTTCCCGTTATAGGTGCGCAAGCTGTCCGGGATGTACCATGTTCAGAACTGCTCCA 121
DB 907 CTTCCCGTTATAGGTGCGCAAGCTGTCCGGGATATACCATGTTCAGAACTGCTCCA 966

QY 122 ACTCAAGCATTTGTATGAGCGAGCGGACATGATGATGATGATGATGATGATGATGATGAT 181
DB 967 ACTCAAGCATTTGTATGAGCGAGCGGACATGATGATGATGATGATGATGATGATGATGAT 1026

QY 182 GCGTTCGGGAGAACAACTCTTCCCGTGTGCGGTAGCGCTCACCCCGCTCGCAGCTA 241
DB 1027 GCGTTCGGGAGAACAACTCTTCCCGTGTGCGGTAGCGCTCACTCCACGCTCGCGCA 1086

QY 242 GGAACGCCAGCTCCCGACACGACAAATACGACGCGCATGCTGCTGTTGGGCGG 301
DB 1087 GGAATGCCAGCTCCCGACATAGCACAATACGACGCGCATGCTGCTGTTGGGCGG 1146

QY 302 CTGCTCTCTGTTCCGCTATGTAGTGGGGATCTCTCGGATCTGTTCTCTGCTCCC 361
DB 1147 CTGCTCTCTGTTCCGCTATGTAGTGGGGATCTCTCGGATCTGTTCTCTGCTCCC 1206

QY 362 AGCTGTTCCACCATCTCGCTTCGGCGCATGAGACGCTGCGAGTCTGCAATGTCTCAATCT 421
DB 1207 AGCTGTTCCACCATCTCTCGCTTCGGCGCATGAGACGCTGCGAGTCTGCAATGTCTCAATCT 1266

QY 422 ATCCCGGCCACATAACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGATGAT 481
DB 1267 ATCCCGGCCATTTATCAGGTACCGCATGCTTGGGATATGATGATGATGATGATGATGATGAT 1326

QY 482 CAAAGCGCTTGTGTATGCGAGTCTCGGATCCCAAGCTGTCGTCGACATGGTGG 541
DB 1327 CAAAGCGCTTGTGTATGCGAGTCTCGGATCCCAAGCTGTCGTCGACATGGTGG 1386

QY 542 CGGGGGCCCATTTGGGGAGTCTTGGGGGCTTGGGGCTTGGCTTCTTCCATTTGGGAACTGGG 601
DB 1387 CGGGGGCCCATTTGGGGAGTCTTGGGGGCTTGGGGCTTGGCTTCTTCCATTTGGGAACTGGG 1446

QY 602 CTAAGGTTTGTATGTGCTACTCTTTGC 632
DB 602 CTAAGGTTTGTATGTGCTACTCTTTGC 632

Dd 1447 CTAAGGTCCTGATTTGGCGCCTACTCTTCGC 1477

RESULT 14
AAQ63753
ID AAQ63753 standard; cDNA to mRNA; 2540 BP.
XX AAQ63753;
XX
XX 30-JAN-1995 (first entry)
DE NANBHV genomic fragment #2.
XX
XX Polymerase chain reaction; PCR; primer; amplif; detection; NANBHV;
KW non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.
OS Synthetic.
XX
XX JF06125777-A.
PN
XX 10-MAY-1994.
PD
XX 20-JUN-1991; 91JP-0247120.
PF
XX 20-JUN-1991; 91JP-0247120.
PR
XX (NAKA/) NAKAMURA T.
PA
XX WPI; 1994-187937/23.
DR
XX
XX The sequences given in AAQ63752-53 represent fragments of the non-A,
PT non-B hepatitis virus (NANBHV) genome. These fragments were amplified
PT using the primers given in AAQ63732-51. These primers were used in the
PT detection of NANBHV. The primers are based on the 5'-terminal region and
PT the core protein coding region. The method allows highly sensitive
XX detection of NANBHV.

PS Disclosure; Page 24-25; 25pp; Japanese.
XX
CC The sequences given in AAQ63752-53 represent fragments of the non-A,
CC non-B hepatitis virus (NANBHV) genome. These fragments were amplified
CC using the primers given in AAQ63732-51. These primers were used in the
CC detection of NANBHV. The primers are based on the 5'-terminal region and
CC the core protein coding region. The method allows highly sensitive
CC detection of NANBHV.

SQ Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 other;

Query Match 88.1%; Score 565.4; DB 15; Length 2540;
Best Local Similarity 93.5%; Pred No. 2.4e-152;
Matches 590; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 2 TGCCCGTTTGTCTTCTCTATCTTCCTCTTGCTTTACTGTCTGTGACCATTCAG 61
| | | | |
Dd 847 TGCCCGTTTGTCTTCTCTATCTTCCTCTTGCTTTGCTGTCTGTGACCATCCAG 906
| | | | |

Qy 62 CTTCCTGGTTATAGGTGCGCAACGTGTCCGGATGTACCATGTACGAACGACTGTCCA 121
| | | | |
Dd 907 CTTCCTGGTTATAGGTGCGCAACGTGTCCGGATGTACCATGTACGAACGACTGTCCA 966
| | | | |

Qy 122 ACTCAAGCATTTGTATAGGACGCGACATGATCATGCACACCCCGGGTGGTGCCT 181
| | | | |
Dd 967 ACTCAAGCATTTGTATAGGACGCGACATGATCATGCATCTCCCGGGTGGTGCCT 1026
| | | | |

Qy 182 GCGTTCCGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTA 241
| | | | |
Dd 1027 GCGTTCCGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACTCCCACGCTCGGCGCA 1086
| | | | |

Qy 242 GGAAACGCCAGCTCCCCACAACGAANAACAGCCACAGTCGATTGCTCGTTGGGGCGG 301
| | | | |
Dd 1087 GGAAATGCGACGCTCCCCACACAGCAATAACAGCCACAGTCGATTGCTCGTTGGGGCGG 1146
| | | | |

Qy 302 CTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGGGATCTGTCTTCTCTGCTCTCCC 361
| | | | |
Dd 1147 CTGCTTCTTCTCGCTATGTACGTGGGGGATCTCTCGGATCTCTGCTCTCTCTCTCCC 1206
| | | | |

CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is the HCV replicon Con 1, used as a basis for
CC the adaptive mutations of the invention.

XX
XX
SQ Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 other;

Query Match 88.1%; Score 565.4; DB 24; Length 9605;
Best Local Similarity 93.5%; Pred. No. 3.9e-152;
Matches 590; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTCTCTATCTCTCTCTGCTTTACTGCTGCTGCTGACCATTCACG 61
DB 847 TGCCCGGTGCTCTTCTCTATCTCTCTGCTTTGCTGCTGCTGACCATTCACG 906
QY 62 CTTCCGCTTATGAGGTGCGCAACGTGTCCGGGATGTACCATGTACGAGAGCTGCTCCA 121
DB 907 CTTCCGCTTATGAGGTGCGCAACGTGTATCCGAGGTGTACCATGTACGAGAGCTGCTCCA 966
QY 122 ACTCAAGCATGTGTATGAGGACGCGGACATGATCATGCACACCCCGGGTGGTGCCT 181
DB 967 ACGAAGCATGTGTATGAGGACGCGGACATGATCATGCATACCCCGGGTGGTGCCT 1026
QY 182 GCGTTCGGGGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTA 241
DB 1027 GCGTTCGGGGAACAACTCTCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCGCGCCA 1086
QY 242 GGAACGCGAGCGTCCCGACACGACATACGACCGCAGCTGATTTGCTGTTGGGCGG 301
DB 1087 GGAACGCTAGCGTCCCGACACGATACGACCGCATGTGCGATTTGCTGTTGGGCGG 1146
QY 302 CTGCTCTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGCTCTCTCTCTCCC 361
DB 1147 CTGCTCTCTGTTCCGCTATGTACGTGGGATCTCTCGGATCTGTTTCTCTCTGCGCC 1206
QY 362 AGCTGTTCCATCTCGCTCCCGGCATGAGACGGTGAGGACTGCAATTTGCTCAATCT 421
DB 1207 AGCTGTTCCATCTCTCGCTCCCGGCACGACAGTACAGGACTGCAATTTGCTCAATAT 1266
QY 422 ATCCCGGCCACATACAGGTACCGGTATGCGTGGGATGATGATGAATGCTCGCTA 481
DB 1267 ATCCCGGCCACGTCAGGTACCGGTATGCGTGGGATGATGATGAATGCTCGCTA 1326
QY 482 CAAACGCGCTGCTGATCGAGCTGCTCCGGATCCCAACAGCTGTGCGTGAATGTTGG 541
DB 1327 CAGAGCGCTAGTGTATCGAGTTACTCCGGATCCCAACAGCTGTGCGTGAATGTTGG 1386
QY 542 CGGGGCGCCATTGGGGAGTCTGGCGGCGCTCGCTACTATTCCATGTTGGGGAACCTGGG 601
DB 1387 CGGGGCGCCATTGGGGAGTCTAGCGGCGCTTGGCTACTATTCCATGTTGGGGAACCTGGG 1446
QY 602 CTAAGGTTTGTATGTGATGCTACTCTTTGC 632
DB 1447 CTAAGGTTCTGATGTGATGCTACTCTTTGC 1477

Search completed: December 19, 2003, 18:51:00
Job time : 182.486 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1651.58 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATGCCCGTGTCTCTTC.....TACTCTTGTCTCTTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estr:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	117.2	18.3	488	9	AV755731
C 2	92.6	14.4	492	9	AV758366
C 3	41.6	6.5	534	14	CD040840
C 4	40.6	6.3	925	29	CNS0091P
					AL053013 Drosophil

ALIGNMENTS

```

RESULT 1
AV755731/c
LOCUS AV755731 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangliang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922
Email: hanzg@hgsc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
FEATURES

```



```
/tissue type="infected host tissue"
/cell_line="p6497"
/dev_stage="48 hour post infection"
/clone_lib="psbH: Infected hypocotyl soybean host. 48 hrs
post infection"
/note="Vector: pBK-CMV, Site 1: EcoRI, Site 2: XhoI;
USDA-IFAPS: Expression of Phytophthora sojae genes during
infection and propagation."
BASE COUNT 101 a 187 c 159 g 87 t
ORIGIN

Query Match 6.5%; Score 41.6; DB 14; Length 534;
Best Local Similarity 47.1%; Pred. No. 2.5;
Matches 128; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 70 TATGAGTGGCAACGTGTCGGGATGTATACCATGTTCACGAACGACATGCTCCAACTCAAGC 129
|||
Db 200 TAGCGGTGCGGAGATTTACGGTATCCGATGCTTCGCCGGCTTCTACAACTGGACC 259
|||

QY 130 ATTGTGTATGAGCAGCAGCATATGATGACACACCCCGGGTGTGCGCTGCGTTCGG 189
|||
Db 260 TCGATGGACCAAGAGAGGCCCGATCATGCTGACCCCAAGACGGTGGCCAACTTAC 319
|||

QY 190 GAGAACTCTTCCCGCTGCTGGGTAGCGCTACCCCGACGCTCGAGTAGGACGCC 249
|||
Db 320 CACTAGCGGGCACCATTCTCGGCTCGAACCGTGTGCTTCGACGTGACAAATATC 379
|||

QY 250 AGCGTCCCGACCAAGCAATATACGACGCCAGCTGATTTGCTGCTGGGGCGGCTGCTCTC 309
|||
Db 380 AACTTCTGACGACGAACGCGCTCTCGAGGTATACGTATCGCGGTGACGCGACCCAC 439
|||

QY 310 TGTTCGGTATGTACGTGGGGGATCTCTCGG 341
|||
Db 440 CGTCCGCGCAACAAGATCTCGGAGGATGCCG 471
|||

RESULT 4
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL053013
VERSION 1
KEYWORDS fly, genomic survey sequence.
SOURCE AL053013.1 GI:4934461
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Trosophilidae; Drosophila.
Genoscope.
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pictet de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
Location/Qualifiers
1. 925

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR19D16"
/clone_lib="RPI-98"
/note="end : TET3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match 6.3%; Score 40.6; DB 29; Length 925;
Best Local Similarity 15.8%; Pred. No. 5.5;
Matches 69; Conservative 165; Mismatches 198; Indels 4; Gaps 1;

QY 122 ACTCAAGCATTTGTATGAGGAGCGGACATGATGACACCCCGGGTGTGCTGCCCT 181
|||
Db 489 AATANANNNTTATTATTANNNNANANANANNANNNNAGCSMCGCKCGSTTBGS 548
|||

QY 182 GCGTTCGGGAGAACACTCTTCCCGTGTGGTAGCGTCAACCCCGCTCGACGCTA 241
|||
Db 549 TTTTSSSGYKGGSGSGSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSS 608
|||

QY 242 GGAACCCGAGCGTCCCGACACGACAAATACGACGCGTGGATTTGCTGTTGGGCGG 301
|||
Db 609 KSSSTBSGCCSCSSKSVGTSCSS-----SSSCSSSSSTSSSTSSSTSSSSSS 664
|||

QY 302 CTGCTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGCTTCTCTCTCC 361
|||
Db 665 SSSSYTTSKTSAGSGSWSAGGGSGTGTSSSSSSSTSSSSSVSSSGSKSTBSGS 724
|||

QY 362 AGCTGTTCACCATCTCGCTCCCGCATGAGCGGTGAGGACTGCAATTGCTCATCT 421
|||
Db 725 BSSSGSSSSSTSSBBSCTSTSSSSSSSTSSCTCCSCSSSYSSSTSSSTSSSTSG 784
|||

QY 422 ATCCCGGCACATAACAGGTCAACGATGCTGGGATATGATGATGAACGTGCGCTA 481
|||
Db 785 SGSSSSSVGTSSSDSTSTCCSCCCYMCCTCCSTYBMCYTCSTSCGSSSSSGGVTKG 844
|||

QY 482 CAACGCGCTGTGTGTATCGCAGCTCTCGGATCCCAAGCTGTGTGACATGTTG 541
|||
Db 845 CGCGSSSTNGMTSSACSSSSSCSSSVSSSSKSSSSSVSSSGSGSVSSSSSSAS 904
|||

QY 542 CGGGGCGCCATGGGG 557
|||
Db 905 KSSSGSGSVSSSGSGS 920
|||

RESULT 5
CNS0091P 590 bp mRNA linear EST 24-NOV-2002
LOCUS wlm1.pk0005.c9 wlm1 Triticum aestivum cDNA clone wlm1.pk0005.c9 5'
DEFINITION end, mRNA sequence.
ACCESSION CA659369
VERSION 1
KEYWORDS Triticum aestivum (bread wheat)
SOURCE Triticum aestivum
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 590)
Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,
Miao, G., Caraher, N. and Hanafey, M.K.
DuPont Wheat cDNA Sequence
Unpublished
Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
Location/Qualifiers
1. 590
```



```
source
1. .590
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Stephens"
/db_xref="taxon:4565"
/clone="wlm1.pk0005.c9"
/tissue_type="leaf"
/notes="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
XhoI; Wheat (Triticum aestivum L.) seedlings 1 hr after
inoculation with Erysiphe graminis f. sp tritici"
BASE COUNT      106 a      218 c      159 g      88 t      19 others
ORIGIN

Query Match      6.3%; Score 40.4; DB 14; Length 590;
Best Local Similarity 47.3%; Pred. No. 5.3;
Matches 122; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 147 GGACATGATCATCACACCCCGGTGGTGGTTCGGGAGAACAACTCTTCGCG 206
Db 43 GTAGATGAGCATGGCGGCGACGCGGCGACGGCTGGCCCTCCACCTCTCTCTC 102
QY 207 CTGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAAGCGCAGCGTCCCGACACGAC 266
Db 103 CGTCTCGCCCTTGCACCGCGCCACACGCGCACCGCGCCCTGCGCTCGGCTCGG 162
QY 267 AATACGACGCGCATGTTGCTGTTGGGGCGGTGCTCTCTGTTTCGCTATGTACGT 326
Db 163 CTTCTTGGCGGCGGCTCACCTCGCGCGCCCGCCACAGACGCGCCCTCCCGCTCCGT 222
QY 327 GGGGATCTCTGGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 386
Db 223 CTCGCGTGGCGGCGCTTGGACGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 282
QY 387 GCATGACGCGTGACGGA 404
Db 283 CAAGTCAAGTGAACGGA 300

RESULT 6
CNS01213      645 bp      DNA      linear      GSS 26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence 17 end of BAC
DEFINITION      BACN08C07 of DrosBAC library from Drosophila melanogaster (fruit
fly); genomic survey sequence.
ACCESSION      AL101589
VERSION      AL101589.1      GI:5613200
KEYWORDS      GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 645)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source
1. .645
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN08C07"

/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/notes="end : T7"
BASE COUNT      28 a      26 c      85 g      92 t      414 others
ORIGIN

Query Match      6.3%; Score 40.4; DB 29; Length 645;
Best Local Similarity 10.9%; Pred. No. 5.5;
Matches 28; Conservative 85; Mismatches 144; Indels 0; Gaps 0;

QY 351 CTTCTCTCCAGCTGTTTCCATCTCCCTCGCGGCGATGAGCGGTGACAGTGTGCA 410
Db 191 SCCGCTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 250
QY 411 TTGCTCAATCTATCCCGGCACATAACAGGTACCGTATGCTTGGGATATGATGAA 470
Db 251 GGSSNNSSNNNGSSSSSSNTSSNNSSNNSSNNSSNNSSNNSSNNSSNNSSNNSS 310
QY 471 CTGGTCTGCTACACGGCCCTGTTGTTATCGAGCTGCTCGGATCCCAAGTGTGCT 530
Db 311 TTNTSSNTSSNNSSATGSSSSSSSGTTTBSGSSSSSSSSSSNNNNNNNNSSNN 370
QY 531 GGACATGTTGCGGGGCGCCATTTGGGAGTCTCTGGGCGCTCGCTACTTCCATGCT 590
Db 371 SNSATTTSSNGSSSSSSSSSSSTSTSTSTSSNNSSNTSTSTSSSTTNTSTSS 430
QY 591 GGGGAACCTGGGCTAAGG 607
Db 431 SSNSTGSSSTTTTNS 447

RESULT 7
CNS01213      525 bp      mRNA      linear      EST 14-JUL-2000
LOCUS      894043G08.y1 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION      Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      BE337089
VERSION      BE337089
KEYWORDS      EST.
SOURCE      Chlamydomonas reinhardtii
ORGANISM      Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 525)
Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
Unpublished
Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu.
Location/Qualifiers
1. .525
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
```


ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 99 a 181 c 160 g 85 t
ORIGIN

Query Match 6.2%; Score 40; DB 10; Length 525;
Best Local Similarity 47.0%; Pred. No. 6.5;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 108 GAACGACTGCTCAACTCAAGCATTTGTATGAGCGGAGCATCATGACACACCCC 167
Db 43 GCACCGCTTCAACACCCACCGTGTGAATTCGCCCGCTCAAGTACCTGTGTCAT 102
QY 168 CGGGTGGTGGCTGCTGCTGGGAGAACAACTTCCCGCTGTGGGTAGCGCTCACCCC 227
Db 103 GGAGGACAAGACCTGCACCTGGAGTCTCGAGTACTGGACCTGGGCAACCTGTCCAA 162
QY 228 CACGCTCGGAGTAGGAAGCGCAGCTGCCACACAGCAATACGACGCCACGTCGATTT 287
Db 163 CGCGCTCAAAAACAACATCTTCATGATCCCAACCCCGTCATCGCGCGCGCGCGCGC 222
QY 288 GCTCGTGGGGCGGCTGCTCTGTTCCGCTATGATAGTGGGGATCTCTGCGGATCTGT 347
Db 223 GGGCGAGCGGCGCGCGCGAGAGTAGCGGAGCGCGCGCGAGCCATGAAGGT 282
QY 348 CTTCTCTGCTCCAGCTGTTTAC 371
Db 283 CAACATGCGCACCTGCTCTAC 306

RESULT 8
BI723733 671 bp mRNA linear EST 19-SEP-2001
LOCUS 1031067P08.y1 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BI723733
VERSION BI723733.1 GI:15699428
KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
REFERENCE 1 (bases 1 to 671)
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.

TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031

JOURNAL Unpublished
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
source Location/Qualifiers
1. .671
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant

Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 139 a 224 c 209 g 99 t
ORIGIN

Query Match 6.2%; Score 40; DB 12; Length 671;
Best Local Similarity 47.0%; Pred. No. 7.1;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 108 GAACGACTGCTCAACTCAAGCATTTGTATGAGCGGAGCATCATGACACACCCC 167
Db 353 GCACCGCTTCAACACCCACCGTGTGAATTCGCCCGCTCAAGTACCTGTGTCAT 412
QY 168 CGGGTGGTGGCTGCTGCTGGGAGAACAACTTCCCGCTGTGGGTAGCGCTCACCCC 227
Db 413 GGAGGACAAGACCTGCACCTGGAGTCTCGAGTACTGGACCTGGGCAACCTGTCCAA 472
QY 228 CACGCTCGGAGTAGGAAGCGCAGCTGCCACACAGCAATACGACGCCACGTCGATTT 287
Db 473 CGCGCTCAAAAACAACATCTTCATGATCCCAACCCCGTCATCGCGCGCGCGCGC 532
QY 288 GCTCGTGGGGCGGCTGCTCTGTTCCGCTATGATAGTGGGGATCTCTGCGGATCTGT 347
Db 533 GGGCGAGCGGCGCGCGGAGAGTAGCGGAGCGCGCGCGAGCCATGAAGGT 592
QY 348 CTTCTCTGCTCCAGCTGTTTAC 371
Db 593 CAACATGCGCACCTGCTCTAC 616

RESULT 9
BI536071 740 bp mRNA linear EST 09-AUG-2002
LOCUS BI536071 MF01SSB cDNA Oryzias latipes cDNA clone MF01SSB002D03 3',
DEFINITION mRNA sequence.

ACCESSION BI536071
VERSION BI536071.1 GI:22194883
KEYWORDS EST.

SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
REFERENCE 1 (bases 1 to 740)
AUTHORS Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.

TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source Location/Qualifiers
1. .740
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="MF01SSB002D03"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/db_stage="segmentation stage 20 - 25"

Db	645	TSNTTTTBTBTSSSTSS
----	-----	---

] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 66 a 162 c 101 g 94 t
ORIGIN

Query Match 6.2%; Score 39.6; DB 9; Length 423;
Best Local Similarity 47.9%; Pred. No. 7.8;
Matches 114; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 195 CAACCTCTCCCGCTGCTGGTAGCGCTCACCCACGCTCGAGCTAGGAGCGCCAGCT 254
Db 83 CATGTCGTCTGCTGCGCGGAGCCACCGCCCTCCGAGACAGCGACGCTCCCTTACGA 142
QY 255 CCCCACACGACAAATACGACGCCACGTCGATTGCTCGTTGGCGCGCTCTCTGTTC 314
Db 143 CCCTAGCGCTCGGCCCTCGCGGCCCTATCTCTCTTCTGCTCTGCTCCCTACTTCTC 202
QY 315 CCGTATGATGTTGGGGATCTCTGCGGATCTGTCTTCTCGTCTCCAGCTGTTTACCAT 374
Db 203 CATCAGGAGACGCGTGACTTCAGCGAGTCCCGCGAGCACCTGGCTAGACAGTTAACAA 262
QY 375 CTCGCTCCCGCGCATGACGCGTGCAGACTGCAATTGCTCAATCTATCCCGGCCAC 432
Db 263 CACGTCTTCCAGCCTGAGCCAGCGCAGTTTGGGAAGGGGCTTCTGCGCCCCCCCAC 320

RESULT 12
CA816001
LOCUS
DEFINITION
CDNA clone CA12E13011VF_E04 5', mRNA sequence.
CA816001
VERSION
KEYWORDS
SOURCE
ORGANISM
Vitis vinifera
Vitis vinifera

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Goes da Silva, F., Lim, H., Iandolino, A., Baek, J., Jones, K., Walker, M.A. and Cook, D.R.
Transcriptional responses of Vitis vinifera to infection by the bacterial pathogen Xylella fastidiosa
Unpublished
Contact: Doug Cook
CAES Genome Facility
UC Davis Department of Plant Pathology
1 Shields Ave., Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcooke@ucdavis.edu

Seq primer: GTATCATGCGACGGTACC.
Location/Qualifiers
1. .621
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA12E13011VF_E04"
/sex="hermaphrodite"
/dev_stage="Mid-season leaf material"
/lab_host="DH5alpha"
/note="Organ: Leaf; Vector: pDNR; Site 1: Sfil; Site 2: Sfil; CA12E1 is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were verified to be infected with the bacterial pathogen, Xylella fastidiosa,

based on a diagnostic assay using PCR and Xylella-specific primer pairs. The plants were asymptomatic at the time of collection, but later developed symptoms. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAGCAGCAGTGGCATACGCGGG-3' and
5'-ATTCTAGAGCGCGCGCGACATG-TT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT 146 a 152 c 143 g 180 t
ORIGIN

Query Match 6.1%; Score 39; DB 14; Length 621;
Best Local Similarity 50.3%; Pred. No. 13;
Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 111 CGACTGCTCCAACTCAAGCATTTGTATGAGCGAGGACATGATCATGCACACCCCGG 170
Db 397 CGACTTGTCTCTATGCAAGCGTAAAGCCAGCAGGAGCTCCGAGCTCTCCGAG 456
QY 171 GTGCGTGCCTCGTTTCGGGAGAACAACTCTTCCGCTCTGCGGTAGCGTCAACCCCGC 230
Db 457 GTGCTCGAGATGCTTGTGTGCTTCTCTCTCTCTAGTCAGGCTCTGCTCATCTGCTA 516
QY 231 GCTCGCAGCTAGAACGCCAGCGTCCCAACAGCAATACGACGCCACGTCGATTGCT 290
Db 517 ACACCTTCGGTGGTTTCGGCAGCGTAGCCGCTGACACCATCGGAAGCATAGCTGCTCGG 576
QY 291 CGTTGGGCGG 301
Db 577 AGTAGTGGG 587

RESULT 13
BI723734
LOCUS
DEFINITION
624 bp mRNA linear EST 19-SEP-2001
1031067F08.y2 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
BI723734
VERSION
KEYWORDS
SOURCE
ORGANISM
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1031
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

Location/Qualifiers
1. .624
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; Stress condition II library, constructed by John Davis and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant

Mon Dec 22 13:28:49 2003

in a medium with bubbling air containing 5% carbon dioxide"

BASE COUNT 70 a 144 c 122 g 63 t

ORIGIN

Query Match 6.0%; Score 38.6; DB 9; Length 399;
Best Local Similarity 50.8%; Pred. No. 14;
Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 30 CTTGGCTTTACTGTCTGTCACCATTCACATTCAGCTTCCGCTTATGAGTGCGCAAGTGTC 89

DB 44 CTTGACCCCTGGACGGCTCGTCCATCGTCGACCACTTCGCGGTGGGACCATCTT 103

QY 90 CGGATGTACCATGTTCACGAACGACTGTCTCAAGCATTTGTATGAGGCGGGA 149

DB 104 CGCGGTGTGCGCGGCGGCAAGCTGACCAATCATCCGCGCGGAGGTGGTGGCGG 163

QY 150 CATGATCATCACACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 209

DB 164 CATGGGATCTACGCTCCCGGACCGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 223

QY 210 C 210

DB 224 C 224

RESULT 15

AV637507

LOCUS

DEFINITION

AV637507

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source

1. .434

/organism="Chlamydomonas reinhardtii"

/mol_type="mRNA"

/strain="C9"

/db_xref="taxon:3055"

/clone="HC074a01_r"

/note="Vector: pBlueScriptII SK-; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"

BASE COUNT 80 a 149 c 130 g 75 t

ORIGIN

Query Match 6.0%; Score 38.6; DB 9; Length 434;

Best Local Similarity 50.8%; Pred. No. 14;

Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 30 CTTGGCTTTACTGTCTGTCACCATTCACATTCAGCTTCCGCTTATGAGTGCGCAAGTGTC 89

DB 21 CTTGACCCCTGGACGGCTCGTCCATCGTCGACCACTTCGCGGTGGGACCATCTT 80

Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). polyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBlueScript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 132 a 210 c 190 g 91 t 1 others

ORIGIN

Query Match 6.1%; Score 39; DB 12; Length 624;

Best Local Similarity 46.6%; Pred. No. 13;

Matches 123; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 108 GAACGACTGCTCAACATCAAGCATTTGTATGAGGCGGACATGATCATGCACACCC 167

DB 353 GCACCGCTTCAACAAACCCACCGTGTGAATTCGCCGCTCAAGTACCTGCTGCCAT 412

QY 168 CGGCTGGCTGCCCTGCTCGGAGAACAACTTCTCCGCTGTGGTACGCTCACCCC 227

DB 413 GGAGGACAGACCTGCGACCTGCGAGTATCGAGTCTGCGACCTGNGCAACCTGTCCAA 472

QY 228 CACGCTGCAGCTAGGAACGCCAGCTCCCGACCAATAGCAGCGCACGTCGATTT 287

DB 473 CGCGCTCAAAACAAACATCTTCATGATCCCAACCGCTCATCGCGGCGCGGGCGC 532

QY 288 GTCCTGTTGGCGGCTGCTCTCTTCCGCTATGATGATGGGGATCTCTGCGGATCTGT 347

DB 533 GGGGACCGCGCGCGCGGAGAGTAGCGGAGCGCGGCGGAGCCATGAAGGT 592

QY 348 CTTCTCTGCTCCGAGCTGTTCCAC 371

DB 593 CAACATGCGCACCTGCTGCTCCAC 616

RESULT 14

AV638521

LOCUS

DEFINITION

AV638521

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source

1. .399

/organism="Chlamydomonas reinhardtii"

/mol_type="mRNA"

/strain="C9"

/db_xref="taxon:3055"

/clone="HC087407_r"

/note="Vector: pBlueScriptII SK-; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"


```

QY      90 CCGGATGTACCATGTACGACGACTGCTCCAACTCAAGCAATTGTGTATGAGGAGCGGA 149
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
81 CCGCGTGTGGCCCGGGACAAAGCTGACCAACATCACCGCGCGGAGCAGGTGGCTGCCGG 140
QY      150 CATGATCATGCACACCCCGGGTGGCGCTGCCCTGGTTCCGGAGAACAACTCTTCCCGCTG 209
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
141 CATGGGCACTACGGTCCCGGACCGGTGTTCTGCATTGCCCTGAAGACGCCCCCGGCTG 200
QY      210 C 210
Db      |
201 C 201

```

Search completed: December 20, 2003, 06:54:34
 Job time : 1658.58 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 46.3008 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-3
Perfect score: 642
Sequence: 1 ATGCCGGTGTCTTCTC.....TACTCTTGTCTCTAATAG 642

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139556

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642	100.0	642	3	US-08-612-973-3
2	642	100.0	642	3	US-08-927-597-3
3	628.2	97.9	795	3	US-08-612-973-5
4	628.2	97.9	795	3	US-08-927-597-5
5	619.8	96.5	2082	3	US-08-612-973-47
6	619.8	96.5	2082	3	US-08-927-597-47
7	619.8	96.5	2433	3	US-08-612-973-49
8	619.8	96.5	2433	3	US-08-927-597-49
9	563.8	87.8	1539	2	US-08-470-426B-17
10	563.8	87.8	1863	2	US-08-470-426B-14
11	559	87.1	742	1	US-08-081-072-18
12	559	87.1	742	1	US-08-449-093A-18
13	559	87.1	932	1	US-08-081-072-15
14	559	87.1	932	1	US-08-449-093A-15
15	555.8	86.6	2116	3	US-08-191-160-21
16	555.8	86.6	9595	3	US-09-014-416-4
17	555.8	86.6	9599	3	US-09-014-416-6
18	544.6	84.8	9472	4	US-08-150-204B-96
19	540.8	84.2	1167	1	US-08-324-977-9
20	540.8	84.2	1167	2	US-08-384-616-9
21	540.8	84.2	1167	2	US-08-304-686A-9
22	540.8	84.2	1167	3	US-09-315-850-9
23	540.8	84.2	1499	1	US-08-324-977-3
24	540.8	84.2	1499	2	US-08-384-616-3
25	540.8	84.2	1499	2	US-08-904-686A-3
26	540.8	84.2	1499	3	US-08-315-850-3
27	540.8	84.2	6039	1	US-08-324-977-11

ALIGNMENTS

RESULT 1
US-08-612-973-3
; Sequence 3, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; NAME/KEY: CDS
; LOCATION: 1..639
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..636

not do MS posted

28 540.8 84.2 6039 2 US-08-384-616-11 Sequence 11, Appl
29 540.8 84.2 6039 2 US-08-904-686A-11 Sequence 11, Appl
30 540.8 84.2 6039 3 US-09-315-850-11 Sequence 11, Appl
31 540.8 84.2 9030 1 US-08-324-977-13 Sequence 13, Appl
32 540.8 84.2 9030 2 US-08-384-616-13 Sequence 13, Appl
33 540.8 84.2 9030 3 US-08-904-686A-13 Sequence 13, Appl
34 540.8 84.2 9030 3 US-09-315-850-13 Sequence 13, Appl
35 540.8 84.2 9416 1 US-08-324-977-1 Sequence 1, Appl
36 540.8 84.2 9416 2 US-08-384-616-1 Sequence 1, Appl
37 540.8 84.2 9416 2 US-08-904-686A-1 Sequence 1, Appl
38 540.8 84.2 9416 3 US-09-315-850-1 Sequence 27, Appl
39 540.8 84.2 9416 4 US-08-823-895A-27 Sequence 21, Appl
40 511.8 79.7 576 1 US-08-086-428B-21 Sequence 21, Appl
41 511.8 79.7 576 2 US-08-468-570-21 Sequence 21, Appl
42 511.8 79.7 576 2 US-08-290-665A-21 Sequence 21, Appl
43 511.8 79.7 576 4 US-08-466-601A-21 Sequence 21, Appl
44 511.8 79.7 576 5 PCT-US95-10398-21 Sequence 21, Appl
45 510.2 79.5 576 1 US-08-086-428B-17 Sequence 17, Appl

RESULT 2

QY 541 GCGGGGCCCCATTGGGAGTCTGGGGGCTCGCCCTACTATTCCATGTTGGGAACTGG 600
Db 541 GCGGGGCCCCATTGGGAGTCTGGGGGCTCGCCCTACTATTCCATGTTGGGAACTGG 600
QY 601 GCTAAGGTTTGTATGATGCTACTCTTTGCTCTCTTAATAG 642
Db 601 GCTAAGGTTTGTATGATGCTACTCTTTGCTCTCTTAATAG 642

RESULT 3
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789
; US-08-612-973-5

Query Match 97.9%; Score 628.2; DB 3; Length 795;
Best Local Similarity 98.8%; Pred. No. 9.7e-156;
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 TGCCCGGTGCTCTTCTCTATCTTCTCTTGGCTTACTGTCTGTGACCAATCCAG 61
Db 155 TGCCCGGTGCTCTTCTCTATCTTCTCTTGGCTTACTGTCTGTGACCAATCCAG 214
QY 62 CTTCCCGCTTATAGGTGCGCAAGTGTCCGGATGTACCATGTCCAGCAAGCTGTCCA 121
Db 215 CTTCCCGCTTATAGGTGCGCAAGTGTCCGGATGTACCATGTCCAGCAAGCTGTCCA 274

QY 122 ACTCAAGCATTTGTATGAGCGAGCGACATGATCATGCAACCCCGGGTGGTGCCT 181
Db 275 ACTCAAGCATTTGTATGAGCGAGCGACATGATCATGCAACCCCGGGTGGTGCCT 334
QY 182 GCGTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGACGCTA 241
Db 335 GCGTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGACGCTA 394
QY 242 GGAACGCCAGCGTCCCCACCAACATACAGCGCCACGCTCGATTTTCTCTGTTGGGGCGG 301
Db 395 GGAACGCCAGCGTCCCCACCAACATACAGCGCCACGCTCGATTTTCTCTGTTGGGGCGG 454
QY 302 CTGCTCTCTGTTCCGCTATGATGTTGGGGATCTCTGCGGATCTGTTCTGCTCTCC 361
Db 455 CTGCTTTCTGTTCCGCTATGATGTTGGGGATCTCTGCGGATCTGTTCTGCTCTCC 514
QY 362 AGCTGTTCCACATCTCGCTCGCGGCATGAGCGGTGAGAGGCTGCAATTGCTCAATCT 421
Db 515 AGCTGTTCCACATCTCGCTCGCGGCATGAGCGGTGAGAGGCTGCAATTGCTCAATCT 574
QY 422 ATCCCGGCCACATAACAGGTTCACCGTATGCTTGGGATATGATGATGAACTGGTGCCTA 481
Db 575 ATCCCGGCCACATAACAGGTTCACCGTATGCTTGGGATATGATGATGAACTGGTGCCTA 634
QY 482 CAACGGCCCTGTTGGTATGCTGCTCGGATGCCCAAGCTGCTGGAACATGTTGG 541
Db 635 CAACGGCCCTGTTGGTATGCTGCTCGGATGCCCAAGCTGCTGGAACATGTTGG 694
QY 542 CCGGGGCCCATTTGGGAGTCTCGCGGCTCGCTACTATTCCATGTTGGGAACTGGG 601
Db 695 CCGGGGCCCATTTGGGAGTCTCGCGGCTCGCTACTATTCCATGTTGGGAACTGGG 754
QY 602 CTAAGGTTTGTATGATGCTACTCTTTGCTCTCTAATAG 642
Db 755 CTAAGGTTTGTATGATGCTACTCTTTGCTCTCTAATAG 795

RESULT 4
US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:

not dp

TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789
US-08-927-597-5

97 q%: Score 628.2; DB 3; Length 795;

C

b
455 CTGCTTTCTGTTCCGCTATGTACGTGGGGACCTCTGGGATCTGCTTCC

515 AGCTGTTACCCATCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCAATT

RESULT 5
US-08-612-973-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
GENERAL INFORMATION:

[illegible]

[illegible]

RESULT 9
 US-08-470-426B-17
 ; Sequence 17, Application US/08470426B
 ; Patent No. 5856458
 ; GENERAL INFORMATION:
 ; APPLICANT: Okamoto, Hiroaki
 ; APPLICANT: Nakamura, Tetsuo
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
 ; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B,
 ; TITLE OF INVENTION: HEPATITIS VIRUS
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,
 ; ADDRESSEE: L.L.P.
 ; STREET: 1850 M Street, N.W., Suite 800
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/470,426B
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-153402
 ; FILING DATE: 12-JUN-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weillacher, Robert G.

:	REGISTRATION NUMBER:	20,531
:	REFERENCE/DOCKET NUMBER:	06/59-47083.1
:	TELECOMMUNICATION INFORMATION:	
:	TELEPHONE:	(202) 659-2811
:	TELEFAX:	(202) 659-1462
:	INFORMATION FOR SEQ ID NO:	17:
:	SEQUENCE CHARACTERISTICS:	
:	LENGTH:	1539 base pairs
:	TYPE:	nucleic acid
:	STRANDEDNESS:	single
:	TOPOLOGY:	unknown
:	MOLECULE TYPE:	DNA (genomic)
:	US-08-470-426B-17	

Query Match	87.8%	Score 563.8;	DB 2;	Length 1539;
Best Local Similarity	93.3%;	Prsd. No. 8.7e-139;		
Matches 589;	Conservative 0;	Mismatches 42;	Indels 0;	Gaps 0;

Qy	2	TGCCCGGTTGCCTTTCCTATCTCCTCTTGCGCTTTACTGTCTGTGACCAATCCAG	61
Dd	506	TGCCCGGTTGCCTTTCCTATCTCCTCTTGCGCTTTACTGTCTGTGACCAATCCAG	565
Qy	62	CTTCCCGCTTATGAGGTGGCAAAGTGTCCGGGATGTACCATTGTACAGAAAGACTGCTCCA	121
Dd	566	CTTCCCGCTTATGAAGTGGCAAGTGTCCGGGATATACCATGTACAGAAAGACTGCTCCA	625
Qy	122	ACTCAAGCATTTGTATGAGCAGCGGACATGATCATGCACACCCCGGTTGCGTGCCT	181
Dd	626	ACTCAAGCATTTGTATGAGCAGCGGACATGATCATGCATCTCCCGGTTGCGTGCCT	685
Qy	182	GGTTTCGGAGAACACTCTTCCCGTGTCTGGGTAGCGCTCACCCCCACGCTGCGAGCTA	241
Dd	686	GGTTTCGGAGGACAACAGCTCCCGTTGTCTGGGTAGCGCTCACTCCCAAGCTGCGGCCA	745
Qy	242	GGAAAGCCAGGTCCTCCACACAGCAATAAGACGCCAGCTGCATTCCTGTTGGGGGG	301
Dd	746	GGAAATGCCAGGGTCCCACATACGACAAATACGACGCCACGTCGACTTGTCTGTTGGGGGG	805
Qy	302	CTGCTCTCTGTCGGCTATGTAGTGGGGATCTTCGCGGATCTGCTCTCTGCTCTCCC	361
Dd	806	CTGCTTCTGCTCGCTATGTAGTGGGGGATCTTCGCGGATCTGCTCTCTCTCCC	865
Qy	362	AGCTGTTTACCATTCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCT	421
Dd	866	AGCTGTTTACCATTCTCGCCTCGCCGGCATGAGACAGTGCAGGACTGCAATTGCTCAATCT	925
Qy	422	ATCCGGCCACATAACAGGTACCGTATGCGTTCGGGATATGATGTAAGACTGTCGCGCTA	481
Dd	926	ATCCGGCCCAATTATCAGGTCACCGCATGCGCTTGGGATATGATGTAAGACTGTCGCGCTA	985
Qy	482	CACAGGCCCTGGTGGTATCCAGTGTCTCGGATCCCCACAGCTGTCTGGACATGTGTGG	541
Dd	986	CACAGGCCCTTAGTGGTGTGTCAGTGTCTCGGATCCCCACAGCTGTCTGGACATGTGTGG	1045
Qy	542	CGGGGGCCCAATTGGGAGTCTCGCGGGCCTCGCTACTATTCCATGCTGGGGAACCTGGG	601
Dd	1046	CGGGGGCCCACTGGGAGTCTCGCGGGCCTTCGCTACTATTCCATGCTAGGGAACCTGGG	1105
Qy	602	CTAAGTTTTGATTTGATGCTACTCTTTGCG	632
Dd	1106	CTAAGTTCCTGATTTGGCGCTACTCTTCGC	1136

RESULT 10
US-08-470-426B-14
; Sequence 14, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE
; TITLE OF INVENTION: APPLICATION FOR
; TITLE OF INVENTION: HEPATITIS VIRUS

Query Match	87.1%;	Score 559;	DB 1;	Length 742;
Best Local Similarity	92.9%;	Prod. No. 1.3e-137;	Indels	0;
Matches	586;	Conservative	Mismatches 45;	Gaps 0
Qy	2	TGCCCGGTTGCTCTTTCTCTAATCTTCCTCTTGCGCTTTACTGTCTGTCTGACATTCGAG	61	
Db	94	TGCCCGGTTGCTCTTTCTCTAATCTTCCTCTTGCGCTTTGCTGTCTGTCTGTGACCATCCGAG	153	
Qy	62	CTTCCCGCTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTCAAGAACGACTGCTCCA	121	
Db	154	CTTCCCGCTTATGAGGTGCGCAACGTTATCCGGATATACCATGTCAAGAACGACTGCTCCA	213	
Qy	122	ACTCAAAGCATTTGTATATGAGGCAGCGGAATGATCATATGCACACCCCGGGTGGGTGCCCT	181	
Db	214	ACTCAAAGTATTGTATATGAGGCAGCGGATGATCATATGCCCGGGTGGGTGCCCT	273	
Qy	182	GCCTTCGGGAGAACAACTCTCCCGCTGTCTGGGTAGCGGTCAACCCCGACGCTCGCAGCTA	241	
Db	274	GCCTTCGGGAGAACAACTCTCCCGTGTCTGGGACGGGTCACTCCACGGTTAGCGGCCA	333	
Qy	242	GGAAACGCCAGCGTCCCGACACAGCAATATACGACCCACGTCGATTTGCTGTTGGGCGG	301	
Db	334	GGAAACACACGTCGCCACATACGACAAATACGACGGATATGCTGCTTTGGGCGG	393	

Query Match	87.9%	Score 563.8	DB 2	Length 1863
Best Local Similarity	93.3%	Pred. No. 9.1e-139		
Matches 589	Conservative 0	Mismatches 42	Indels 0	Gaps 0
QY	2	TGCCGGTTGCTCTTTCTCTATCTCTCTTTGCTTTACTGCTCTCTGACCATCCAG	51	
DB	830	TGCCGGTTGCTCTTTCTCTATCTCTCTTTGCTTTGCTCTCTTTGACCATCCAG	889	
QY	62	CTTCCGGTTATAGGTCCGCAACGTGTCCGGGATGTACCATGTACAGACGCTGTCCA	121	
DB	890	CTTCCGGTTATAGGTCCGCAACGTGTCCGGGATATACCATGTACAGACGCTGTCCA	949	
QY	122	ACTCAAGCATTTGTATAGGACGCGGACATGATCATGCACACCCCGGTGCGTGCCCT	181	
DB	950	ACTCAAGCATTTGTGTATAGGACGCGGACATGATCATGCATCTCCCGGGTGCGTGCCCT	1009	
QY	182	CGCTTCGGGAGAACAACTCTTCCGCTGTCTGGGTAGCGCTACCCCCACGCTCGCAGCTA	241	
DB	1010	CGCTTCGGGAGCAACAACGCTCCCGTTGCTGGGTAGCGCTCACTCCACACCTCGCGGCCA	1069	
QY	242	GGACGCCAGCGTCCCCACACGACAATACGACGCCACGCTCGATTTGCTCGTTGGGGCGG	301	
DB	1070	GGAAATGCCAGGGTCCCCACTACGACAATACGACGCCACGCTCGACTTGTCTGTTGGGGCGG	1129	
QY	302	CTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTCGGATCTGTCTTCTCGTCTCCC	361	
DB	1130	CTGCTTTCTGCTCCGCTATGTACGTGGGGGATCTCTCGGATCTGTCTTCTCGTCTCCC	1189	
QY	362	AGCTTTCCACATCTCGGCTCGCCGGCATGACGGTGCAGGCTGCAATTGCTCAATCT	421	
DB	1190	AGCTTTCCACCTTCTCGGCTCGCCGGCATGACGAGTGCAGGCTGCAATCTCAATCT	1249	
QY	422	ATCCCGGGCCACATAACAGGTCAACCGTATGGCTTGGGATATGATGATGAACGTGGCGCTA	481	
DB	1250	ATCCCGGGCATTTATACAGGTCAACCGATGGCTTGGGATATGATGATGAACGTGGCACCTA	1309	

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 3010.09 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-5
Perfect score: 795
Sequence: 1 AGTTGGTAAGTCAATCGA.....TACTTTGCTCCTTAATAG 795

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	795	100.0	795	6	A48667	A48667 Sequence 5
2	795	100.0	795	6	AR157325	AR157325 Sequence
3	795	100.0	795	6	AX452754	AX452754 Sequence
4	795	100.0	795	6	AX685006	AX685006 Sequence
5	778.8	98.0	2082	6	A48709	A48709 Sequence 47
6	778.8	98.0	2082	6	AR157350	AR157350 Sequence
7	778.8	98.0	2082	6	AX452796	AX452796 Sequence
8	778.8	98.0	2082	6	AX685048	AX685048 Sequence
9	778.8	98.0	2433	6	A48711	A48711 Sequence 49
10	778.8	98.0	2433	6	AR157351	AR157351 Sequence
11	778.8	98.0	2433	6	AX452798	AX452798 Sequence
12	778.8	98.0	2433	6	AX685050	AX685050 Sequence
13	714.8	89.9	9379	14	AF207766	AF207766 Hepatitis
14	711.6	89.5	9379	14	HCVPOLYP	AJ000009 Hepatitis
15	711	89.4	9418	14	HCV132996	AJ132996 Hepatitis
16	710	89.3	9386	14	AF165056	AF165056 Hepatitis
17	708.4	89.1	9386	14	AF165055	AF165055 Hepatitis
18	708.4	89.1	9410	14	HPCK1R2	D50481 Hepatitis C
19	706.8	88.9	3296	14	AB008441	AB008441 Hepatitis
20	706.8	88.9	3296	14	AB008443	AB008443 Hepatitis
21	706.8	88.9	9379	14	AF165052	AF165052 Hepatitis
22	706.8	88.9	9379	14	AF207761	AF207761 Hepatitis
23	705.2	88.7	3296	14	AB008442	AB008442 Hepatitis
24	705.2	88.7	3296	14	AB008446	AB008446 Hepatitis
25	705.2	88.7	9379	14	AF165051	AF165051 Hepatitis
26	705.2	88.7	9598	14	AB049101	AB049101 Hepatitis
27	703.6	88.5	3296	14	AB008447	AB008447 Hepatitis
28	703.6	88.5	9547	14	AB049091	AB049091 Hepatitis
29	703.6	88.5	9605	6	AX739971	AX739971 Sequence
30	703.6	88.5	9605	14	HCJ238799	AJ238799 Hepatitis
31	703.6	88.5	11076	6	AX036252	AX036252 Sequence
32	703.6	88.5	11076	6	AX036258	AX036258 Sequence
33	703.6	88.5	11076	6	AX036260	AX036260 Sequence
34	703.6	88.5	11076	6	AX036262	AX036262 Sequence
35	702	88.3	1880	14	HPC5TRJ4	D00832 Hepatitis C
36	702	88.3	2540	6	E04260	E04260 cDNA encodi
37	702	88.3	2540	6	E04805	E04805 cDNA to 5'
38	702	88.3	2540	6	E07391	E07391 cDNA encodi
39	702	88.3	3296	14	AB008445	AB008445 Hepatitis
40	702	88.3	9033	14	HCJ238800	AJ238800 Hepatitis
41	702	88.3	9379	14	AF165048	AF165048 Hepatitis
42	702	88.3	9448	14	HPCJ483	D13558 Hepatitis C
43	700.4	88.1	1587	14	HPCCP8P1	D10687 Hepatitis C
44	700.4	88.1	1615	14	HPCNS1SPJ	M74813 Hepatitis C
45	700.4	88.1	9344	14	AB049096	AB049096 Hepatitis

ALIGNMENTS

RESULT 1
A48667
LOCUS A48667 Sequence 5 from Patent WO9604385.
DEFINITION A48667
ACCESSION A48667
VERSION A48667.1 GI:2302380
KEYWORDS
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 795)
AUTHORS Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 5 15-FEB-1996;

Mon Dec 22 13:28:51 2003

Db 661 CTCGGGATCCCAAGCTGTCTGTGACATGTGTGGGGGCCCAATTGGGAGTCTCTGGCG 720

QY 721 GGTCTCGCTACTATTCATGTGGGAACTGGGCTTAAGTTTTCATTGTGATGCTACTC 780

Db 721 GGTCTCGCTACTATTCATGTGGGAACTGGGCTTAAGTTTTCATTGTGATGCTACTC 780

QY 781 TTTGCTCCCTAATAG 795

Db 781 TTTGCTCCCTAATAG 795

RESULT 2

ARL57325

LOCUS ARL57325 795 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 5 from patent US 6245503.

ACCESSION ARL57325

VERSION ARL57325.1 GI:16218258

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 795)

AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.

TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use

JOURNAL Patent: US 6245503-A 5 12-JUN-2001;

FEATURES Location/Qualifiers

source 1..795

BASE COUNT 130 a 240 c 231 g 194 t

ORIGIN

Query Match 100.0%; Score 795; DB 6; Length 795;

Best Local Similarity 100.0%; Pred. No. 3.2e-167; Indels 0; Gaps 0;

Matches 795; Conservative 0; Mismatches 0;

QY 1 ATGTTGGGTAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60

QY 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCGAGGCCCTGGGCGATGGCGTCCGGTT 120

Db 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCGAGGCCCTGGGCGATGGCGTCCGGTT 120

QY 121 CTGGAGGACGGCGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTCTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTCTCTATCTTC 180

QY 181 CTCTGGCTTCTGTCCTGTCTGACCGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCT 240

Db 181 CTCTGGCTTCTGTCCTGTCTGACCGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCT 240

QY 241 TCCGGGATGATACATGTACGACGACCTGCTCCAACTCAAGCATTTGTATGAGGCAGCG 300

Db 241 TCCGGGATGATACATGTACGACGACCTGCTCCAACTCAAGCATTTGTATGAGGCAGCG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCCGC 360

Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCCGC 420

Db 361 TGCTGGGTAGCGCTACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCCGC 420

QY 421 ATAGAGCCGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

Db 421 ATAGAGCCGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

QY 481 GGGGACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

Db 481 GGGGACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

QY 541 CATGAGACGGTGCAGGACTGCAATTCATCTATCCCGGCCACATACCGGTACCGGT 600

INNOGENETICS NV (BE)

Other publication CA 2172273 960215

Other publication AU 3382495 960304.

Location/Qualifiers

1..795

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

1..792

/notes="unnamed protein product"

/codon_start=1

/protein_id="CAA03129.1"

/db_xref="GI:2302381"

/translation="MLGKVIDTLTCGFADLVGIVPLVGLPGLGAARALAHGVRVLEDD

VNATGLPCGSPSIFLLALLSLTPASAYERNVSGMYHTNDCSNSIVYEADAM

IMHTPGCVPCVRNNSRWCVALPTLAARNASVPTTIRRHVDLLVGNAAFCASMYV

GDLCGSVFLVSQFTTSPRRHETVQDNCISYPGHITGRMAWDMMNWSPTTALVVS

QLLRIPQAVDMVAGAHGVLGLAYYSVMGNWAKVLIVMLLFAP"

mat_peptide 1..789

/product="unnamed"

BASE COUNT 130 a 240 c 231 g 194 t

ORIGIN

Query Match 100.0%; Score 795; DB 6; Length 795;

Best Local Similarity 100.0%; Pred. No. 3.2e-167; Indels 0; Gaps 0;

Matches 795; Conservative 0; Mismatches 0;

QY 1 ATGTTGGGTAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60

QY 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCGAGGCCCTGGGCGATGGCGTCCGGTT 120

Db 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCGAGGCCCTGGGCGATGGCGTCCGGTT 120

QY 121 CTGGAGGACGGCGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTCTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTCTCTATCTTC 180

QY 181 CTCTGGCTTCTGTCCTGTCTGACCGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCT 240

Db 181 CTCTGGCTTCTGTCCTGTCTGACCGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCT 240

QY 241 TCCGGGATGATACATGTACGACGACCTGCTCCAACTCAAGCATTTGTATGAGGCAGCG 300

Db 241 TCCGGGATGATACATGTACGACGACCTGCTCCAACTCAAGCATTTGTATGAGGCAGCG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCCGC 360

Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCCGC 420

Db 361 TGCTGGGTAGCGCTACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCCGC 420

QY 421 ATAGAGCCGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

Db 421 ATAGAGCCGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

QY 481 GGGGACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

Db 481 GGGGACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

QY 541 CATGAGACGGTGCAGGACTGCAATTCATCTATCCCGGCCACATACCGGTACCGGT 600

Db 541 CATGAGACGGTGCAGGACTGCAATTCATCTATCCCGGCCACATACCGGTACCGGT 600

QY 601 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660

Db 601 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660

QY 661 CTCGGGATCCCAAGAGTGTCTGTGACATGTGGGGGGGCCCAATTGGGAGTCTCTGGCG 720

Db 541 CATGAGCGGTGACGAGCTGCAATTCCTCAATCTATCCCGGCGACATTAACGGGTACCGT 600
Qy 601 ATGGCTTGGGATATGATGATGAATCGTCCGCTACAAACCGCCCTGGTGGTATGCGAGCTG 660
Db 601 ATGGCTTGGGATATGATGATGAATCGTCCGCTACAAACCGCCCTGGTGGTATGCGAGCTG 660
Qy 661 CTCGGATCCCAACGCTGCTGGACATGCTGGCGGGGCGCCATTTGGGGAGTCTTGGCG 720
Db 661 CTCGGATCCCAACGCTGCTGGACATGCTGGCGGGGCGCCATTTGGGGAGTCTTGGCG 720
Qy 721 GGCTCGCTTACTTATCCATGCTGGGAACTGGGCTAAGGTTTGGTGGTGGTGGTGGT 780
Db 721 GGCTCGCTTACTTATCCATGCTGGGAACTGGGCTAAGGTTTGGTGGTGGTGGTGGT 780
Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795

RESULT 3
AX452754
LOCUS
DEFINITION
Sequence 5 from Patent EP1211315.
AX452754
ACCESSION
AX452754.1 GI:21712439
VERSION
KEYWORDS
SOURCE
Hepatitis C virus
ORGANISM
Hepatitis C virus
Viruses; serNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
1
Maertens, G., Bosman, F., de Martynoff, G. and Buysse, M.A.
TITLE
Recombinant vectors for producing hcv envelope proteins
JOURNAL
Patent: EP 1211315-A 5 05-JUN-2002
Innogenetics N.V. (BE)

FEATURES
source
Location/Qualifiers
1..795
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..792
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38067.1"
/db_xref="GI:21712440"
/translations="MLGKVIDTLTCGPADLVGYIPLVGAFLGGAARALAHGVRLVDG
VNYATGNLPGCSFSIFLLALLSCLTYPASAYEVRNVSVMYHVTNDCSNSSIVYEAADM
IMHTGCPVPCVRNNSRCWVALTPILARNASVPTTIHRHVDLLVGAAPFCSAMTV
GDLGCVFLVQLFTSPRRHETVQDCNCSYIPGHITGHRMAMWMMNPPPTALVVS
QLLRIPQAVVDMVAGAHGVLGLAYYSVMGNWAKVLIVMLLEAP"
mat_peptide
1..789
/product="unnamed"
BASE COUNT 130 a 240 c 231 g 194 t
ORIGIN

Query Match 100.0%; Score 795; DB 6; Length 795;
Best Local Similarity 100.0%; Pred. No. 3.2e-167;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
Qy 61 CCCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTGGCGATGCGTCCGGTT 120
Db 61 CCCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTGGCGATGCGTCCGGTT 120
Qy 121 CTGAGGACCGGTGAATATGCAACAGGAAATTTGCCCGTTCCTCTTCTATCTTC 180
Db 121 CTGAGGACCGGTGAATATGCAACAGGAAATTTGCCCGTTCCTCTTCTATCTTC 180
Qy 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGCACAGCTG 240

Db 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGCACAGCTG 240
Qy 241 TCCGGGATGTACCATGTACGAAACGAGTGTCTCCAACTCAAGCATTTGTGTATGAGGACG 300
Db 241 TCCGGGATGTACCATGTACGAAACGAGTGTCTCCAACTCAAGCATTTGTGTATGAGGACG 300
Qy 301 GACATGATCATGACACCCCGGGTGGTCCCTGCTGCGGTTCGGGAGAAACAATCTTCCCG 360
Db 301 GACATGATCATGACACCCCGGGTGGTCCCTGCTGCGGTTCGGGAGAAACAATCTTCCCG 360
Qy 361 TGCTGGGTAGCGCTCACCCCGGCTGCGAGCTAGGAACGCCAGCGTCCCGACACACA 420
Db 361 TGCTGGGTAGCGCTCACCCCGGCTGCGAGCTAGGAACGCCAGCGTCCCGACACACA 420
Qy 421 ATACGACGCCACGCTCGATTTGCTGCTTGGGGCGGCTGCTTCTGCTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGCTCGATTTGCTGCTTGGGGCGGCTGCTTCTGCTTCCGCTATGTACGTG 480
Qy 481 GGGGACCTCTGCGGATCTGCTTCTCCGCTGCCAGCTGTTCACCATCTCGCCTCGCGG 540
Db 481 GGGGACCTCTGCGGATCTGCTTCTCCGCTGCCAGCTGTTCACCATCTCGCCTCGCGG 540
Qy 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATTAACGGGTCA 600
Db 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATTAACGGGTCA 600
Qy 601 ATGGCTTGGGATATGATGATGAATCGTCCGCTACAAACGCCCGCTGGTGGTATCGCAGCTG 660
Db 601 ATGGCTTGGGATATGATGATGAATCGTCCGCTACAAACGCCCGCTGGTGGTATCGCAGCTG 660
Qy 661 CTCGGATCCCAACGCTGCTGGACATGCTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 720
Db 661 CTCGGATCCCAACGCTGCTGGACATGCTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 720
Qy 721 GGCTCGCTTACTTATCCATGCTGGGAACTGGGCTAAGGTTTGGTGGTGGTGGTGGT 780
Db 721 GGCTCGCTTACTTATCCATGCTGGGAACTGGGCTAAGGTTTGGTGGTGGTGGTGGT 780
Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795

RESULT 4
AX685006
LOCUS
DEFINITION
Sequence 5 from Patent WO0205548.
AX685006
ACCESSION
AX685006.1 GI:29371411
VERSION
KEYWORDS
SOURCE
Hepatitis C virus
ORGANISM
Hepatitis C virus
Viruses; serNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
1
Maertens, G., Bosman, F. and Buysse, M.A.
TITLE
Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL
Patent: WO 0205548-A 5 18-JUL-2002;
INNOGENETICS N.V. (BE)

FEATURES
source
Location/Qualifiers
1..795
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..792
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD86520.1"
/db_xref="GI:29371412"
/translations="MLGKVIDTLTCGPADLVGYIPLVGAFLGGAARALAHGVRLVDG
VNYATGNLPGCSFSIFLLALLSCLTYPASAYEVRNVSVMYHVTNDCSNSSIVYEAADM

mat_peptide		130 a	240 c	231 g	194 t	/product="unnamed"	
<p>BASE COUNT 130 a 240 c 231 g 194 t</p> <p>ORIGIN</p> <p>Query Match 100.0%; Score 795; DB 6; Length 795;</p> <p>Best Local Similarity 100.0%; Pred. No. 3.2e-167;</p> <p>Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>							
Qy	1	ATGTTGGTAAAGTCATCGATACCC	TACATCGCGGCTTCGCGACCT	CTGTGGGTACATT	60		
Db	1	ATGTTGGTAAAGTCATCGATACCC	TACATCGCGGCTTCGCGACCT	CTGTGGGTACATT	60		
Qy	61	CCGCTCGTCGGCGCCCTTAGGGGGCGCT	GCAGGGCCCTGCAGCATCGCGGTTCGGGTT	120			
Db	61	CCGCTCGTCGGCGCCCTTAGGGGGCGCT	GCAGGGCCCTGCAGGGCCCTGCAGCATCGCGGTTCGGGTT	120			
Qy	121	CTGAGGACGCGTAACTATGCAACAGGGAA	TTTGCCCGGTTGCTTTTCTCTATCTTC	180			
Db	121	CTGAGGACGCGGTAACATATGCAACAGGGAA	TTTGCCCGGTTGCTTTTCTCTCTATCTTC	180			
Qy	181	CTCTTGGCTTTGCTGTCTGTCAACCGTTC	CAGCTTCCGCTTATGAAGTGCACACGTG	240			
Db	181	CTCTTGGCTTTGCTGTCTGTCAACCGTTC	CAGCTTCCGCTTATGAAGTGCACACGTG	240			
Qy	241	TCCGGGATGACCATGTACAGAAAGACGTCT	CCAACTCAAGCAATGTGTATGAGGCAGCG	300			
Db	241	TCCGGGATGACCATGTACAGAAAGACGTCT	CCAACTCAAGCAATGTGTATGAGGCAGCG	300			
Qy	301	GACATGATCATGCAACCCCGGGTGCCTGCG	TTCCGGTCCGGGAGAACACTCTTCCCGC	360			
Db	301	GACATGATCATGCAACCCCGGGTGCCTGCG	TTCCGGTCCGGGAGAACACTCTTCCCGC	360			
Qy	361	TGCTGGTAGCGCTCACCCCGGTCGAGCTAG	GAAACGCGAGCGTCCCCACACGACA	420			
Db	361	TGCTGGTAGCGCTCACCCCGGTCGAGCTAG	GAAACGCGAGCGTCCCCACACGACA	420			
Qy	421	ATACGACGCGCATCGATTGCTGTTGGGGCG	GCTGCTTTGTTCCGCTATGTACGTG	480			
Db	421	ATACGACGCGCATCGATTGCTGTTGGGGCG	GCTGCTTTGTTCCGCTATGTACGTG	480			
Qy	481	GGGGAACCTTGCGGATCTGTTCTTCGCTCC	AGCTGTTACCATCTCGCCTCGCGCG	540			
Db	481	GGGGAACCTTGCGGATCTGTTCTTCGCTCC	AGCTGTTACCATCTCGCCTCGCGCG	540			
Qy	541	CATGAGACGCTGCAGGACTGCAATTGCTCAA	TATATCCGGGCAATACGCGGTCAACGT	600			
Db	541	CATGAGACGCTGCAGGACTGCAATTGCTCAA	TATATCCGGGCAATACGCGGTCAACGT	600			
Qy	601	ATGGCTTGGATATGATGATGAATCGTGCCT	TACAAACGCGCCCTGGTGGTATCGCAGCTG	660			
Db	601	ATGGCTTGGATATGATGATGAATCGTGCCT	TACAAACGCGCCCTGGTGGTATCGCAGCTG	660			
Qy	661	CTCCGGATCCCAAGCTGTCTGGACATGGTGG	CGGGGCCCAATGGGGAGTCTCGGCG	720			
Db	661	CTCCGGATCCCAAGCTGTCTGGACATGGTGG	CGGGGCCCAATGGGGAGTCTCGGCG	720			
Qy	721	GGTCTCGCTACTATTCCATGGTGGGACCTG	GGCTAAGGTTTTTGATGTGACTACTC	780			
Db	721	GGTCTCGCTACTATTCCATGGTGGGACCTG	GGCTAAGGTTTTTGATGTGACTACTC	780			
Qy	781	TTTGCTCCCTAATAG	795				
Db	781	TTTGCTCCCTAATAG	795				
<p>RESULT 5</p> <p>A48709</p> <p>LOCUS</p> <p>Sequence 47 from Patent WO9604385.</p> <p>denaturation</p> <p>A48709</p> <p>linear</p> <p>DNA</p> <p>2082 bp</p> <p>PAT 07-MAR-1995</p>							


```
Db 364 TGGGTAGCGCTCACCCCGAGCTCGCAGCTAGGAAAGCAGAGCTCCCGACACACAATA 423
Qy 424 CGAGCCACGTCGATTTGCTCGTTGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
Db 424 CGAGCCACGTCGATTTGCTCGTTGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
Qy 484 GACCTCTGGGATCTGCTTCCTCGCTCCAGCTGTTACCACTTCGCTCCGCGGCT 543
Db 484 GACCTCTGGGATCTGCTTCCTCGCTCCAGCTGTTACCACTTCGCTCCGCGGCT 543
Qy 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTCAACCGTATG 603
Db 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTCAACCGTATG 603
Qy 604 GCTTGGATATGATGATGAACTGGTGCCTTACCAAGCGCCCTGCTGATGCGAGTCTC 663
Db 604 GCTTGGATATGATGATGAACTGGTGCCTTACCAAGCGCCCTGCTGATGCGAGTCTC 663
Qy 664 CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCGCCATTTGGGGAGTCTTGGCGGGT 723
Db 664 CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCGCCATTTGGGGAGTCTTGGCGGGT 723
Qy 724 CTGCGCTACTATTCCTGTTGGGAACTGGGCTAAAGTTTGTGATGCTACTCTTT 783
Db 724 CTGCGCTACTATTCCTGTTGGGAACTGGGCTAAAGTTTGTGATGCTACTCTTT 783
Qy 784 GC 785
Db 784 GC 785

RESULT 6
LOCUS AR157350 2082 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 47 from patent US 6245503.
ACCESSION AR157350
VERSION AR157350.1 GI:16218284
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2082)
AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 47 12-JUN-2001;
FEATURES
source Location/Qualifiers
BASE COUNT 366 a 634 c 500 g 482 t
ORIGIN

Query Match 98.0%; Score 778.8; DB 6; Length 2082;
Best Local Similarity 99.7%; Pred. No. 1.3e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TTGGTAAAGTCTATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 63
Db 4 TTGGTAAAGTCTATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 63
Qy 64 CTGCTCGGCGCCCTTACGAGGCGCTGCGAGGCGCTGCGGAGTCTGCGGCTGCGGCTG 123
Db 64 CTGCTCGGCGCCCTTACGAGGCGCTGCGAGGCGCTGCGGAGTCTGCGGCTGCGGCTG 123
Qy 124 GAGACGGCTGACTATGCAACAGGAAATTTGCCCGGTTGCTCTTCTATCTTCTC 183
Db 124 GAGACGGCTGACTATGCAACAGGAAATTTGCCCGGTTGCTCTTCTATCTTCTC 183
Qy 184 TTGGCTTTGCTGCTCTGCTGACCGTTTCCAGCTTTCGCTTATGAAGTGCACAGTCTCC 243
Db 184 TTGGCTTTGCTGCTCTGCTGACCGTTTCCAGCTTTCGCTTATGAAGTGCACAGTCTCC 243
Qy 244 GGGATGTACCATGTGTACGAAAGCATGTCTCCAACTCAAGCATTTGTGTATGAGCGAGCGGAC 303
```

```
Db 244 GGGATGTACCATGTGTACGAAAGCATGTCTCCAACTCAAGCATTTGTGTATGAGCGAGCGGAC 303
Qy 304 ATCATCATGACACCCCGGGTGCCTCGCTTGGGAGAACAACTCTTCCCGCTGC 363
Db 304 ATCATCATGACACCCCGGGTGCCTCGCTTGGGAGAACAACTCTTCCCGCTGC 363
Qy 364 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAAGCAGAGCTCCCGACACACAATA 423
Db 364 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAAGCAGAGCTCCCGACACACAATA 423
Qy 424 CGAGCCACGTCGATTTGCTCGTTGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
Db 424 CGAGCCACGTCGATTTGCTCGTTGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
Qy 484 GACCTCTGGGATCTGCTTCCTCGCTCCAGCTGTTACCACTTCGCTCCGCGGCT 543
Db 484 GACCTCTGGGATCTGCTTCCTCGCTCCAGCTGTTACCACTTCGCTCCGCGGCT 543
Qy 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTCAACCGTATG 603
Db 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTCAACCGTATG 603
Qy 604 GCTTGGATATGATGATGAACTGGTGCCTTACCAAGCGCCCTGCTGATGCGAGTCTC 663
Db 604 GCTTGGATATGATGATGAACTGGTGCCTTACCAAGCGCCCTGCTGATGCGAGTCTC 663
Qy 664 CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCGCCATTTGGGGAGTCTTGGCGGGT 723
Db 664 CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCGCCATTTGGGGAGTCTTGGCGGGT 723
Qy 724 CTGCGCTACTATTCCTGTTGGGAACTGGGCTAAAGTTTGTGATGCTACTCTTT 783
Db 724 CTGCGCTACTATTCCTGTTGGGAACTGGGCTAAAGTTTGTGATGCTACTCTTT 783
Qy 784 GC 785
Db 784 GC 785

RESULT 7
LOCUS AX452796 2082 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 47 from Patent EP1211315.
ACCESSION AX452796
VERSION AX452796.1 GI:21712481
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 47 05-JUN-2002;
FEATURES
source Location/Qualifiers
1. .2082
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
<1. .2079
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD38082.1"
/db_xref="GI:21712482"
CDS
```


AX685048	LOCUS	AX685048	2082 bp	DNA	linear	PAT 29-MAR-2003
SEQUENCE	DEFINITION	Sequence 47 from Patent WO02055548.				
AX685048	ACCESSION	AX685048				
AX685048.1	VERSION	AX685048.1	GI:29371453			
KEYWORDS	KEYWORDS	Hepatitis C virus				
SOURCE	SOURCE	Hepatitis C virus				
ORGANISM	ORGANISM	Hepatitis C virus				
		Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.				
REFERENCE	REFERENCE	1				
AUTHORS	AUTHORS	Maertens, G., Bosman, F. and Buysse, M.A.				
TITLE	TITLE	Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use				
JOURNAL	JOURNAL	Patent: WO 02055548-A 47 18-JUL-2002;				
		INNOGENETICS N.V. (BE)				
FEATURES	FEATURES	Location/Qualifiers				
source	source	1..2082				
		/organism="Hepatitis C virus"				
		/mol_type="genomic DNA"				
		/db_xref="taxon:11103"				
		<1..2079				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="CAD86535.1"				
		/db_xref="GI:29371454"				
		/translation="NLGVKIDTLTCTGFPADLVGYIPLVGAFLPGGAARLAHGVRLVDG				
		VNYATGNLPGCSFISFLALLSCLTVPASAYEVNVSVMYHVTNDCNSISIVVEADM				
		IMHTPGCPVCVRENNSSRWALPTLAARNAASVPTTIRHVDLIVGAFAFCAMTV				
		GDLCGSLVLSQLFTIISPRBETVQDCNSIYPGHTGHRMAMDMNMNPTTALVVS				
		QLLRIQPAVDVMAVAGAHQKILQVNTNHSIIRNLKNCNDSLOTGPFALFYKHKFNSSG				
		SDTLGLVLSFPGSAQKIQVNTNHSIIRNLKNCNDSLOTGPFALFYKHKFNSSG				
		CPEGLASCRSIDKFAQGMGLTYTPNSDSVLIILNTPPRGNWFCCTWMNGTFTKCGP				
		TPSPVVGTTDRFGVPTNMGANDSVLIILNTPPRGNWFCCTWMNGTFTKCGP				
		PCNIGAGNNTLTCTGTCFRKHPATYARCGSPMLTPRCMVHPIFYRLMHPCTVNT				
		IFKVMYGVGEHREFAECNWRGRCDEDRSELSPLLSFTTEWOILPCSTFTTIP				
		ALSTGLIHLHQNIVDVOLYGVSAVSLVIKWEYVLLIFLLADARICACLMWMLLI				
		AQEAALLENLVLNAAVAGAHGTLISFLVFFCAAAYIKGRLLVPGAAVAFYGVWPLLLI				
		LLALPPRAYA"				
mat_peptide	mat_peptide	1..2076				
BASE COUNT	BASE COUNT	366 a 634 c 600 g 482 t				
ORIGIN	ORIGIN	/product="unnamed"				
Query Match	Query Match	98.0%; Score 778.8; DB 6; Length 2082;				
Best Local Similarity	Best Local Similarity	99.7%; Pred. No. 1.3e-163;				
Matches 780; Conservative	Matches 780; Conservative	0; Mismatches 2; Indels 0; Gaps 0;				
Qy	Qy	4 TTGGTGAAGTCAFCGATACCTTTACATGGCGCTTGCCGACCTCGTGGGTACATTCGG 63				
Db	Db	4 TTGGTGAAGTCAFCGATACCTTTACATGGCGCTTGCCGACCTCGTGGGTACATTCGG 63				
Qy	Qy	64 CTGCTCGGCGCCCTAGGGGCGCTGCCAGGCGCTGCGCGCATGGCGTCCGGGTCTG 123				
Db	Db	64 CTGCTCGGCGCCCTAGGGGCGCTGCCAGGCGCTGCGCGCATGGCGTCCGGGTCTG 123				
Qy	Qy	124 GAGACGCGTGAACATATGCAACAGGGAATTCGCCGGTTCCTTTCTATCTTCCTC 183				
Db	Db	124 GAGACGCGTGAACATATGCAACAGGGAATTCGCCGGTTCCTTTCTATCTTCCTC 183				
Qy	Qy	184 TTGGCTTTGTGTCCTGTCGTGACCGTTCAGCTTCGGCTTATGAAGTGCGCAACGTGTC 243				
Db	Db	184 TTGGCTTTGTGTCCTGTCGTGACCGTTCAGCTTCGGCTTATGAAGTGCGCAACGTGTC 243				
Qy	Qy	244 GGGATGTACCATGTGACGAACGACTGCTCCAACCTCAAGCAATTTGTATGAGGACGGGAC 303				
Db	Db	244 GGGATGTACCATGTGACGAACGACTGCTCCAACCTCAAGCAATTTGTATGAGGACGGGAC 303				
Qy	Qy	304 ATGATCATGCACACCCCGGGTGTGCGCTCGGTTCGGGAGACAACTCTTCCCGCTGC 363				
Db	Db	304 ATGATCATGCACACCCCGGGTGTGCGCTCGGTTCGGGAGACAACTCTTCCCGCTGC 363				
Qy	Qy	364 TGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCGAGCGTCCCGACACACAATA 423				
Db	Db	364 TGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCGAGCGTCCCGACACACAATA 423				
Qy	Qy	424 CGACGCCAGCTCGATTTGCTGTGGGGCGCTGCTTTCTGTTCCGCTATGTAGTGGG 483				
Db	Db	424 CGACGCCAGCTCGATTTGCTGTGGGGCGCTGCTTTCTGTTCCGCTATGTAGTGGG 483				
Qy	Qy	484 GACCTCTGGGATCTGCTTCTCTCTGCTCCAGCTGTTCCACCATCTCGCCTCGCGCAT 543				
Db	Db	484 GACCTCTGGGATCTGCTTCTCTCTGCTCCAGCTGTTCCACCATCTCGCCTCGCGCAT 543				
Qy	Qy	544 GAGACGCGTGAAGTCAATTCCTCAATCTATCCCGGCCATACCGGTCACCGTATG 603				
Db	Db	544 GAGACGCGTGAAGTCAATTCCTCAATCTATCCCGGCCATACCGGTCACCGTATG 603				
Qy	Qy	604 GTTGGGATATGATGAACTGCTGCTTACAAACCGGCGCTGCTGATCGAGCTGCTC 663				
Db	Db	604 GTTGGGATATGATGAACTGCTGCTTACAAACCGGCGCTGCTGATCGAGCTGCTC 663				
Qy	Qy	664 CGGATCCCAACAGCTGCTGGGACATGGTGGGGGGGCCCATTTGGGGAGTCTTGGGGGT 723				
Db	Db	664 CGGATCCCAACAGCTGCTGGGACATGGTGGGGGGGCCCATTTGGGGAGTCTTGGGGGT 723				
Qy	Qy	724 CTCGCTACTATTCATGGTGGGAACTGGGCTTAAGGTTTGTATGTGCTACTCTTT 783				
Db	Db	724 CTCGCTACTATTCATGGTGGGAACTGGGCTTAAGGTTTGTATGTGCTACTCTTT 783				
Qy	Qy	784 GC 785				
Db	Db	784 GC 785				

VERSION AR157351.1 GI:16218285
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2433)
AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 49 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..2433
/organism="unknown"
BASE COUNT 434 a 745 c 714 g 540 t
ORIGIN
Query Match 98.0%; Score 778.8; DB 6; Length 2433;
Best Local Similarity 99.7%; Pred. No. 1.3e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 TTGGTAAGTATCGATACCTTACATGCGGCTTCGCCACCTCGTGGGTACATTCCG 63
Db 355 TTGGTAAGTATCGATACCTTACATGCGGCTTCGCCACCTCGTGGGTACATTCCG 414
QY 64 CTCGTGCGCGCCCTAGAGGGGCGCTGCCAGGCGCCCTGGCGCATGGCGTTCGGTCTG 123
Db 415 CTCGTGCGCGCCCTAGAGGGGCGCTGCCAGGCGCCCTGGCGCATGGCGTTCGGTCTG 474
QY 124 GAGGACGCGTGAATATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 183
Db 475 GAGGACGCGTGAATATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTAAAGTGCACAGCTGTC 243
Db 535 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTAAAGTGCACAGCTGTC 594
QY 244 GGGATGTACCATGTCAAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGACGGAC 303
Db 595 GGGATGTACCATGTCAAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGACGGAC 654
QY 304 ATGATCATGACACCCCGGGTGGTCCGCTCGCTTGGGAGAACAACTCTTCCCGTGC 363
Db 655 ATGATCATGACACCCCGGGTGGTCCGCTCGCTTGGGAGAACAACTCTTCCCGTGC 714
QY 364 TGGGTAGCGCTACCCCGCCAGCTCGAGTAGAAGCGCGCTCCCGACCAACGACATA 423
Db 715 TGGGTAGCGCTACCCCGCCAGCTCGAGTAGAAGCGCGCTCCCGACCAACGACATA 774
QY 424 CGACGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 775 CGACGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
QY 484 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
Db 835 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
QY 544 GAGACGCTGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603
Db 895 GAGACGCTGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
Db 955 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
QY 664 CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
Db 1015 CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
QY 724 CTCGCTACTATTCATGCTGGGAACTGGGCTTAAGGTTTGTATGATGCTACTCTTT 783
Db 1075 CTCGCTACTATTCATGCTGGGAACTGGGCTTAAGGTTTGTATGATGCTACTCTTT 1134
QY 784 GC 785

Db 1135 GC 1136
RESULT 11
AX452798 2433 bp DNA linear PAR 06-JUL-2002
LOCUS Sequence 49 from Patent EP1211315.
DEFINITION AX452798
ACCESSION AX452798
KEYWORDS AX452798.1 GI:21712483
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.-A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: Ep 1211315-A 49 05-JUN-2002;
INNOVATIONS N.V. (BE)
FEATURES Location/Qualifiers
source 1..2433
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
CDS 1..2430
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38083.1"
/db_xref="GI:21712484"
/translation="MSTNPFQRTKENTNRPPQDVKFGPGQIIVGGVYLPRRPRRL
GVRATRTKTSERSQPRRRQPIPKARREGRANWAPGYWPLVGNMGWAGWLLSPRG
SRPWSGPTDPRRRNLGKVIDTLTCFADLVGIPLVGAPLGGAPALAHGVRLRD
GVNATGNLPGCSFISILLALLSLTVPASAYEVNRVSGMTHVNDSCNSISIVYEAD
MIMTPGCVPCVRNNSRSCWALTVALAARNSVPTTIRRHVDLLVGLAAAPFALV
VMDLGSVELNSQLFTTSPRHETVQDCNCSIYPGTHGHMMDMMNHSPTSCVAM
SOLLIPQAVDMVAGAHGVLAVYVSNVGNWAKLVNMLLPAGDGHTRVSGAA
ASDTRGLVSLFSPGSAQKIQLVNTNGSHINRTALNCNDSLOTGFFAALFYKHFNSS
GDEPLASCRSIDKFAQGPPLTYENSSDQRPYCWYAPCGIWPASVCGPVYC
FPSPVAVGTTDRFGVPTYNMGANDSVLLIINNIRPRGNWFGCTWNGTGTCTCGG
PCNLGGAGNNTLTCTDPCFHEATYARCGSPWLTFRCHVHPYRLWHVPCVNF
TIFKVMVGVGVEHFEACNWTGRCDLDRSELSPLLSLTETTEQIILCSFTTL
PALSTGLIHLHONIVDVQVLYGVSAVSLVWKVEYVLLFILLADARICACLWMLL
IAQAAEALENLVNAAAVAGAGHTLSFLVFCFAAWYIKGRVLPGAAYAFYGVNPLLL
LLLLUPPRAIA"
mat_peptide 1..2427
/product="unnamed"
BASE COUNT 434 a 745 c 714 g 540 t
ORIGIN
Query Match 98.0%; Score 778.8; DB 6; Length 2433;
Best Local Similarity 99.7%; Pred. No. 1.3e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 TTGGTAAGTATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 63
Db 355 TTGGTAAGTATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 414
QY 64 CTCGTGCGCGCCCTAGAGGGGCGCTGCCAGGCGCCCTGGCGCATGGCGTTCGGTCTG 123
Db 415 CTCGTGCGCGCCCTAGAGGGGCGCTGCCAGGCGCCCTGGCGCATGGCGTTCGGTCTG 474
QY 124 GAGGACGCGTGAATATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 183
Db 475 GAGGACGCGTGAATATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTAAAGTGCACAGCTGTC 243
Db 535 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTAAAGTGCACAGCTGTC 594
QY 244 GGGATGTACCATGTCAAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGACGGAC 303
Db 595 GGGATGTACCATGTCAAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGACGGAC 654


```
QY 304 ATGATCATGACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTCCCGCTGC 363
D 655 ATGATCATGACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTCCCGCTGC 714
QY 364 TGGGTAGCGCTCACCCCGCGCTCGAGCTAGGAACGCCAGCGTCCACACGACAA 423
D 715 TGGGTAGCGCTCACCCCGCGCTCGAGCTAGGAACGCCAGCGTCCACACGACAA 774
QY 424 CGACGCCACGCTGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGACGTGGG 483
D 775 CGACGCCACGCTGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGACGTGGG 834
QY 484 GACCTCTGCGGATCTGCTTCTGCTGCTCCAGCTGTTCCACATCTCGCTCGCGGAT 543
D 835 GACCTCTGCGGATCTGCTTCTGCTGCTCCAGCTGTTCCACATCTCGCTCGCGGAT 894
QY 544 GAGACGCTGACGACTGCAATTTGCTCAATCTATCCCGGCCACATAGCGGTCA 603
D 895 GAGACGCTGACGACTGCAATTTGCTCAATCTATCCCGGCCACATAGCGGTCA 954
QY 604 GCTTGGGATATGATGAATGCTGCTGCTTACCAACGCCCTGCTGCTGATCGCAGCTGCTC 663
D 955 GCTTGGGATATGATGAATGCTGCTGCTTACCAACGCCCTGCTGCTGATCGCAGCTGCTC 1014
QY 664 CGATATCCCAAGCTGCTGCTGACATGCTGCGGGGCCCAATGGGGAGTCTTGGCGGGT 723
D 1015 CGATATCCCAAGCTGCTGCTGACATGCTGCGGGGCCCAATGGGGAGTCTTGGCGGGC 1074
QY 724 CTGCGCTACTATCCATGCTGGGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
D 1075 CTGCGCTACTATCCATGCTGGGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
QY 784 GC 785
D 1135 GC 1136

RESULT 12
LOCUS AX685050
DEFINITION Sequence 49 from Patent WO0205548.
ACCESSION AX685050
VERSION AX685050.1 GI:29371455
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1 Maertens, G., Bosman, F. and Buyse, M.A.
Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 49 18-JUL-2002;
INNOGENETICS N.V. (BE)
FEATURES
Location/Qualifiers
source 1..2433
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..2430
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD86536.1"
/db_xref="GI:29371456"
/translations="MSTNPKPQRKTRKTRNRPDVKFPQGGQIVGGVLLPRRGPRL
GVTRATKTSRSQPRGRQPIPKARPEGRWAQPGYPWPLYNMGWAGWLLSPRG
SRPSWGTDPDRSRNLGKVIDITCGFADLVGYIPLVGAPLGGARALAHGVRLVD
GVNATGNLPGCSFSIFLLALLCLVTPASAYEVNRVSTTHRVHVDLLVGAARFCSAM
MIMHTPCGVCVRENSSRCWALPTPLARNASVPTTIRRHVDLLVGAARFCSAM
VGLCGSVFLVSQFTISPRRHTQDCNSIYPGHTIGHRMAWMMNWSPTALVV
SQLRIPQAVWVAGAHGVLGLAGLAVYVMGNWAKVLVYVMLLFPAGVDGTRVSGGAA
ASDTRGLVSLFSPGSAKIQLVNTNGSWHINRTALNCNDSLQTFGAALFYKHFNS
```

```
mat_peptide 1..2427
BASE COUNT 434 a 745 c 714 g 540 t
ORIGIN
Query Match 98.0%; Score 778.8; DB 6; Length 2433;
Best Local Similarity 99.7%; Pred. No. 1.3e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 TTGGTAAGGTCAATCGATACCTTACATCGGGCTTCCAGCCAGCTCGTGGGGTACATTCG 63
D 355 TTGGTAAGGTCAATCGATACCTTACATCGGGCTTCCAGCCAGCTCGTGGGGTACATTCG 414
QY 64 CTGCTCGGGCCCGCTTAGGGGGCGCTCCAGGGCCCTGGCGCATGGCGTTCGGGTCTG 123
D 415 CTGCTCGGGCCCGCTTAGGGGGCGCTCCAGGGCCCTGGCGCATGGCGTTCGGGTCTG 474
QY 124 GAGACGCGCTGAACATATGCAACAGGGAATTTGCGGGTTCCTCTTCTATCTTCTC 183
D 475 GAGACGCGCTGAACATATGCAACAGGGAATTTGCGGGTTCCTCTTCTATCTTCTC 534
QY 184 TTGGCTTTGCTGCTGCTGCTGCGCTTCAGCTTCGCTTATGAGTGGCGGAGTCTCC 243
D 535 TTGGCTTTGCTGCTGCTGCTGCGCTTCAGCTTCGCTTATGAGTGGCGGAGTCTCC 594
QY 244 GGGATGTACCATGTCTACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGAC 303
D 595 GGGATGTACCATGTCTACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGAC 654
QY 304 ATGATCATGACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTTCCCGCTGC 363
D 655 ATGATCATGACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTTCCCGCTGC 714
QY 364 TGGGTAGCGCTCACCCCGCGCTCGAGCTAGGAACGCCAGCGTCCACACGACAA 423
D 715 TGGGTAGCGCTCACCCCGCGCTCGAGCTAGGAACGCCAGCGTCCACACGACAA 774
QY 424 CGACGCCACGCTGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGACGTGGG 483
D 775 CGACGCCACGCTGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGACGTGGG 834
QY 484 GACCTCTGCGGATCTGCTTCTGCTGCTCCAGCTGTTCCACATCTCGCTCGCGGAT 543
D 835 GACCTCTGCGGATCTGCTTCTGCTGCTCCAGCTGTTCCACATCTCGCTCGCGGAT 894
QY 544 GAGACGCTGACGACTGCAATTTGCTCAATCTATCCCGGCCACATAGCGGTCA 603
D 895 GAGACGCTGACGACTGCAATTTGCTCAATCTATCCCGGCCACATAGCGGTCA 954
QY 604 GCTTGGGATATGATGAATGCTGCTGCTTACCAACGCCCTGCTGCTGATCGCAGCTGCTC 663
D 955 GCTTGGGATATGATGAATGCTGCTGCTTACCAACGCCCTGCTGCTGATCGCAGCTGCTC 1014
QY 664 CGGATCCCAAGCTGCTGCTGAGCATGCTGGCGGGGCCCAATGGGGAGTCTTGGCGGGT 723
D 1015 CGGATCCCAAGCTGCTGCTGAGCATGCTGGCGGGGCCCAATGGGGAGTCTTGGCGGGC 1074
QY 724 CTGCGCTACTATTCATGCTGGGGAACTGGGCTAAGGTTTTGATGTGATGCTACTTT 783
D 1075 CTGCGCTACTATTCATGCTGGGGAACTGGGCTAAGGTTTTGATGTGATGCTACTTT 1134
QY 784 GC 785
D 1135 GC 1136
```


[illegible]


```

Db      1404  CTTGCCCTACTATTCCATGGTGGGGAACCTGGCGTAAGGTCTTGATGTGATGCTACTCTTC 1463
Qy      784  GC 785
Db      .1464  GC 1465

RESULT 14
LOCUS   HCVPOLYP
DEFINITION   Hepatitis C virus complete genome sequence.
ACCESSION   AJ000009
VERSION      AJ000009.1
KEYWORDS    complete genome; core protein; E1 protein; E2 protein; NS2 protein;
            NS3 protein; NS4a protein; NS4b protein; NS5a protein; NS5b
            protein; p7 protein; polyprotein.

SOURCE    Hepatitis C virus
ORGANISM   Hepatitis C virus
           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
           Hepacivirus.
REFERENCE  1 (bases 1 to 9379)
AUTHORS   Trowbridge R. and Gowans E.J.
TITLE     Molecular cloning of an Australian isolate of hepatitis C virus
JOURNAL   Arch. Virol. 143 (3), 501-511 (1998)
MEDLINE   98232263
PUBMED    9572551
REFERENCE  2 (bases 1 to 9379)
AUTHORS   Trowbridge R.
TITLE     Direct Submission
JOURNAL   Submitted (02-JUL-1997) Trowbridge R., Hepatitis Unit, Sir Albert
            Sakzewski Virus Research Centre, Royal Children's Hospital, Herston
            Road, Brisbane, Queensland Q 4029, AUSTRALIA
FEATURES   Location/Qualifiers
            1..9379
            /organism="Hepatitis C virus"
            /virion
            /mol_type="genomic RNA"
            /isolate="HCV-A"
            /db_xref="taxon:11103"
            1..9379
            /notes="genomic RNA for polyprotein precursor, complete
            genome"
            1..324
            /codon_start=1
            /product="polyprotein"
            /protein_id="CAA03854.1"
            /db_xref="GI:2764398"
            /translat="MSTNPKPQKTRKTRNRNRRPQDVKFGGQIIVGGVLLPRRGPRL
            GVRLAKRTSRSOPRGRROPIPKARRPEGRWAQPGYPWPYLGNEGMWAGWLLSPRG
            SRPSMGSPDRPRSRNLGKVIDITLTCGADLMGYIPLVGPALGGAARALAHGVRLSD
            GVNAYAGNLPGCCFSFLLGLLCLITIPASAYEVRNVSGVHVINDCSNASIYERAD
            MIMHVPQCPVPCVNRSSRCWALTPTLAARNSAVPTAIRRHVDLLVGAATCSAMY
            VGLDCSGFVLAQLFTFSRPHETVQDCNSIYPGHITGHRMADMMWNSPTAALV
            SOLLRIPOALVDMVAGHMGVLAGLAYSVMGNKAKVLIWLLFAGVDGHTHTGGVA
            QDRLRTFGFSLGPKIOIOLVNTGSMWIRNLTALNCNDSLTGWLAAALFYTHSPNAS
            GCPERASCHPIDEFAGGPGPIYAEHSSDQRPYCHYAPQPCGIVPASEVCGPYVC
            FTSPVUVGTDRHGVGTYSWGENGTDLVLLNTRPPQGNWFGCTWNGGTGFTKTCGG
            FPCNIGVGNGNTLTCTDCPRKHPATYTKCGSGPMLTPRCMVDYFYRLWHYPTVNF
            TIFKRMVGVGVHRLSAACNWRGERCDLDRSLSPLLLSTTEWQVLPGSFTTL
            PALSTGLIHQNIIDVOYLYGIVSVVFAIKWEYVVLFLLLADARACALWMLL
            IAOEAALENLVNAAASVAGAHGILSFVFCAAWYIKGRVLPGRAYAYGYWPLL
            LLALPPRAVALDREMAASCGAVFVGLVLLTSLSPHYKEFLARLWLYQYITRAEL
            LQWIPPLNVRGDRDAIILITCVHPPELIDITKLILAMIGPPMWLVQAVIKVPYVR
            AQGLIRACMLVAGHYVQMAKLAGLTSYVDHLTPLODWAHGGRLDLAVVE
            PVVFSDMETXIIITWADTAACGDIISGLPVSAARRGRIILGPADSLKEQGWLLAPIT
            AYSGQRTGLLGCIIITSLTRDKNOVEQVQVSTATOSFLATCVNGCVTVVYHAGSK
            TLAGKGPVTQMTVNDQDLVGPAPGAEASLTPTCTGSSDLYLVTRHADYIVRRRG
            DSRGSLSPRPVSYLYKSSGGPLLCPSCSHAVGIFRAAVCTRGVAKAVDFVFPVSMETT
            MRSFPVTDNSSPAVPETFQVAHLHAPTGSKSTKVPAAAYAAQYKVLVLPNSVAATL
            GFGAYMSKAHGTDPNRTITGVRTITTTGAPITYSTYKFLADGGSGGAYDIIICDECHS

```

```

mat_peptide 325..897
mat_peptide 898..1473
mat_peptide 1474..2562
mat_peptide 2563..2751
mat_peptide 2752..3402
mat_peptide 3403..5295
mat_peptide 5296..5457
mat_peptide 5458..6240
mat_peptide 6241..7581
mat_peptide 7582..9354
mat_peptide 9358..9379

```

```

3'UTR 1993 a 2830 c 2694 g 1962 t
ORIGIN

```

```

Query Match      89.5%; Score 711.6; DB 14; Length 9379;
Best Local Similarity 94.4%; Pred. No. 1.3e-148;
Matches 738; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

```

Qy 4 TTGGGTAAGTCTATCGATACCTTTACATGCGGCTTCGCCGACCTCGTGGGGGTACATTCGG 63
Db 679 TTGGGTAAGTCTATCGATACCTTTACATGCGGCTTCGCCGACCTCATGGGGGTACATTCGG 738
Qy 64 CTGTCGGCGCCCGCTAGGGGGCGCTGCCAGGCGCTCGCGGATGCGCGTCCGGTCTG 123
Db 739 CTGTCGGCGCCCGCTTAGGGGGCGCGCCAGGCGCTTCGCCGATGCGCGTCCGGTCTG 798
Qy 124 GAGGACGCGTGAACATATGCAACAGGGAAATTTGCCCGGTTGCTCTTCTCTATCTTCTC 183
Db 799 GAGGACGCGTGAACATATGCAACAGGGAAATTTGCCCGGTTGCTCTTCTCTATCTTCTC 858
Qy 184 TTGGCTTTGCTGCTGCTGCTGCTGACCGTTCAGCTTCCGCTTATGAAGTGCAGCGTCC 243
Db 859 TTGGGTTTGTGCTGCTGCTGCTGACCATCCAGCTTCCGCTTATGAAGTGCAGCGTCC 918

```

```

TDTTILIGITVLDQATAGARLVLATATPPGSVTVPHPNIEVALNSNTGIEPIFYK
AIPETIKGGRHLIFCHSKCDELAAKLSGLNNAVAYRGLDVSIPTSGDVVVA
TDLMTKGTGDFSDVDCNTCVTQDFSLDPTFTIETTPQDVAVSRRGRTRG
RRGIYRFTPGERSGFWSSVLCCEYDAGCAWVELTAPTSVRLRAYLNTPLGPVQ
DHLEFESVFTGLTHDAHFLSOTKQAGENFYLTAQATVCARQAAPPSPDMWKC
LIRLKTPLHGPTLLYRLGAVONEVLTPIIKYIMACMSADLEVTSTVWLVGGVLA
ALAAVCLTTSVIVGRILISGRPALIPREVLQFEDMEECASHUPIYEOGMQLAE
QFKOKALGLQATATQKAAAVPVVESKQALEAFWAKHMWNFISGIIYLAGSLTLPN
LAIASLMAFTASITSLTQHTLLFNILGWAQVAAQLAPPASAFAVGAGIAGAAGSI
GLKGLVDILAGYAGVAGALVAFKVMGSEMPSTEDLVNLLPALISPGALVGVCAA
ILRHVPGSGVQVMNRLIAFASRGNHVSPTHVPESDAAAVTOILSITITOLLYK
ELHWNEDCSTPCSGSLWDWDWICTVLITDFKWLQSKLLPRLPGIPYSCORGKY
GVWRGDMQITPCGAQITGHVNGSMRIVGPKTCSNTWHGTFPINAYITGCTPSP
APNSYRALMRVAAEYEVTRVGDPHVTGTTNDNVKPCQCPAPEFTTVDGVRLLR
YAPACPLRDEVTFOVLYVVSQQLPCEPEPDDVVVVTSMITDPSHITAEAKRRL
DRGSPSLASSASQASPSLCACTTRHDSPADLIEANLWROEMGNITRVESEN
KVILDSFDPLRAEEDEREVSAAETILKRSRRPRAMPVWARPDYNPPLAESWKDPDY
VPPVHGCELPPTKAPPIPPRKRRTVLTSTVSSALAEATKTFGSSSSSVDSTG
ATAPPOPPNDITGSDVSCSMPLLEGEPDPLSDSGSWSTVSEASBDVVCCSMS
YTTGALITPCAABESKLPINALNSLLRHHNNVYATTSASQKQKVTDFRLQVLD
DHYRDLKEMKAKASTVKAKLSEVEACKLTPPHSARSKFGYGAQKVRNLSKAVNH
VSTLPOAVMGPSYGFQSPQORVEFLVNAWKSCKCEMGAYDTRCFDSVTESDIRV
ESISYOCCLDAPARQAISLTERLYIGGLTNSKGNCCGYRRCRAGSGVLTSCGNTL
TCYLKASAAAKLQDCTMLVCGDDLVVICSAGTQEDASLURVFTFTEATRYSPAG
DPQPEYDLELITSCSNVSAHDASGRVYVLTDRPTTFLARAAMETARTHTPVNSWL
GNIMYAPTILWARMILMTHFPIILAAEQLEKALDQIYGACYSIEPLDLPQIQRHL
GLSFLSHSYSGEINRVASCLRLKGVPLRARWHRARSVRALKLSGGRAATCGRYL
FNWAVTKLKLTPIPAASQLDLSNWFVAGVSGGDIYHLSRARPRPMWCLLLLSVGV
CIVLLPNR"

```


		/protein_id="CAB41950.1"		/db_xref="GI:4753719"	
		/translation="MSTNPKPORKTKRNTNRRPODKVFPGGGOIVGGVYLLPRRGRPL GYRATKTSERSQPRGRKQPIPKARPEGRWAQCYGWPPLYGNEGLGAGWLLSPRG SRPMPGTPRRRSRLKGVLDITLTCGFADLMGYIPLVVCAPLGABARALAHGVYLED GUNYATGNLPGCSFSIFLLALLSLCTIPASAYEVRNVSGVHVHTVDCNSNLSIVYTAD MIMHTGCVPCVREDNSSRCVALPTLAARNGSVPTTAIRRHVDLLVGAALFCSAM VGFQCSGVLSQFLTSLPRRHETVOECNSIYPGHVTHGRMAWMMNMSPTTALVV SOLLIPQAVDNVAGAHGVLGAYLVSMVGNWAKLLIMLLFAGVDDGGGTRTIGG SQQAAGLSVMSFVSFGSKILQINTNGSWHNRTALNCNDSLNTGFLAALFYAHKFN SSCQPERMASCRPIDRPAQGWGPIITTAEPGSLDQRPYCWYAPQPCGIVPASEVCGPV YCTPTSPVVGTTDRSGVPTQGDNETDVLNNTRPQGNWFGCTWNNTGFTKTC GTPPCNIGGLNNTLICTPTDCPKHSEATYTKGSGPWLTPRCIVDYPRLRHHYPCVY NFTLFEKVMYVGGIEHRLAACNWTGRCNEDRSELSPLUUSTTETWQILPCSFPT TLTAPSTGLHLHQNIVDQVLYIGISVFAIKWEYILLLFLADARACALMMW LLLAQAALENLVNMAAGAHGILGFLVFCFAAYIKGRVLPFGAAYALYSVWPL AHLQWTFPLNVRGRDAIILLTCAVHPELIPDITKLLAILGLPMLVQAGIIRAPYF VRAQGLIRACMLRVKAGHYVQMAFMKLAALGTGYVYDHLTPLRMAHAGRLDAVA VEPVFSDMETKII TWGADTAACGDIILGLPVSARRGKETILGPASLSGQWRLLAP ITAYSQTRGLGCIITSLTRDRKNQVGEVQVSTATQOSFLATCVNGVCWYTHGAG SKTLAGKGPITQMYTNVDQDLVGMPPAPSGARSLTFTCGSSDLXILVSRHADVLPVR RGDSRGLSPRUSVLKSGSGGPLLCPSHGVGVI FRAAVCTRGVAKADVFPVBSME TTTSPVFTDNSSPPAVPQFVAHLHAPHTSGSKSTKVPAAVAAQGYKVLVLPNSVAA TLSFGAYMSKAGVDPISRTIGVRTITGAPITTYGKFLADGCGSGAYDIITICBC HSTDSTILGIGTLDQETAGAKLVVIAATATPPGSVTVVPHNIEEVALNTEGIEIPY GKAIPETIRGRHLIFCHSKKCDLAAQLSGLGINAVAYIRGLDVSVIPTSQDVVG VATDALMTGTFDPSVIDNCTVCTQVDFSLDPTFTIETITVPQDAVSRSQRGRTG RGRGIVRFVTPGPRSGMFDSSVLCYCDAGCAWYELTPTAETSVRLURALINTFGLPV CODHLEFWESVFTGLTHIDAHFTLSQKAGNFPFLVAYQATVCARAQAQPPSWDQW KCLIRKPTLHMGPTLILYRLGAVONEVTLTHPTKYIMAGSADLEVTVLTVLVGGV LAAALAYCLITGTSVIVGRIITLSGKPAIPDRGVLYREDEMBECASHLPVIEQMOL AEQFKALGALLOTATQAAAAAPVBSKRALEAFWAKHWSPIGIGIYLAGLSTLP GNPAIASLMAFTASVTSPLTHTLILNGWVAAQLAPPSAASA FVGAAGAAVVC STGLKVLVDILAGVAGALVAFKVMGEMPESTEDLVNLLPAILSPGALVGVVC AAILRRHVGRGEGAVQWNRILIAFASRGNHVSPTHYPVPSDAAARVTQILSSLTQL LRLHQWLNEDCSTPCSGWLRDWDNICTVLTDFTKWLQSKLLPRLFPFVFFSCRG YKVMRGDGIQMTTCCPQCAQITGVHVGKSMRI VGPRTCSNTHGTFPFINATYGPCTP SPAPYFRALMRVAAEYIVRVGDVHYVTGMTDNVCKPCQVPAPPEFFVEVGVRL HRYAPACKPFLREEVTFQGLNQYLVSQLPCEPEPGVYVLTSMITDPSHITAEAKR RLDRGPPSLASSSASQSAFSLKATCTTRHSDPADLIGIENALLWRQBMGNLIRVES ENKVVILDSFDPLRAEBEREISVPAEILRRSRKFPRAPIWARPDYNPPLLESKDP DYVPVHVGCPLPAKAPVPVPPRRKTVLTSTVSSALAEALATKTGSSSESAVDS GTATAPGPGSDGDTGSDAESCSMPPLEGGPDPLSDGSSWSTVSEASEDVVCCS MSTYTGALITPCAAEESKLPINALSNLHNNVYATTSRSASOROKVTYFDRLOV LDHYVMDLKEMKAKASTVKARLLSVEACKLTPHSAKSKFGYKADKVRNLSRAIN HINSVMEDLLEDTETPIDITIMAKNEVFCVQPEKGRKPARLIVPDLGVRVCEKMAL RVGESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNCNGCRRCRASGVLTTCGN TLCYLKASAAKRAAKLODCTMLVCGDDLVV CESAGTOEDAAASLRVFEAMTRISAP PGDPPOEYDLELITSCSNVSAHDASGRKRYVYLTDPPTTLARAAMETABTPVNS WLGNIIMYAPTLWARNVLMHFPFSILLAOLEKALDCQIYGACYSIEPLDLPQIIQR HLGLSPALHSYSPGEINRVASCLRGLVPLRVWRHARSVRAKLLSRRGAATCGK YLFNWAVRTKLKLTPIPAASARLDLSGMFVAGVSGGDIYHLSLSPARPMPFWMCLLLSLV GVGIYLLPNR"			
		BASE COUNT 1861 a 2843 c 2716 g 1998 t			
		ORIGIN			
		Query Match 89.4%; Score 711; DB 14; Length 9418;			
		Best Local Similarity 94.3%; Pred. No. 1.7e-148;			
		Matches 738; Conservative 0; Mismatches 45; Indels 0; Gaps 0;			
QY	4	TTGGTGAAGTTCATCGATACCCCTTACATGCGGCTTCGCCACCTCGTGGGTACATTCG 63			
Db	695	TTGGTGAAGTTCATCGATACCCCTCACAATGCGGCTTCGCCACCTCATGGGGTACATTCG 754			
QY	64	CTCTGTCGGGCCCCCTAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123			
Db	755	CTCTGTCGGGCCCCCTTAGGGGGCTGCCAGGGCCCTTCGGCGCATGGCGTCCGGGTTCTG 814			
QY	124	GAGGACGGCTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTCTC 183			
Db	815	GAGGACGGCTGAATATGCAACAGGGAATCTGCCCGGTTGCTCTTTCTATCTTCTC 874			

QY	244	GGGATGTACCATGTACAGAACGACTGCTCCAATCAAGCATTTGTGTATGAGCGCGGAC 303	
Db	919	GGGTGTACCATGTACAGAACGACTGCTCCAACGCAAGCATTTGTGTATGAGCGCGGAC 978	
QY	304	ATGATCATCACACCCCGGGTGGCTGCTCGGTTTCGGGAGAAACAATCTTTCCTCGCTGC 363	
Db	979	ATGATCATGACAGTCCCGGGTGGCTGCTCGGTTTCGGGAGCAACAATCTTTCCTCGCTGC 1038	
QY	364	TGGGTAGCGCTCAACCCCAAGCTCCGACGTAGGAAACGCCAGGCTCCCAACACGACAATA 423	
Db	1039	TGGGTAGCGCTCAACCCCAAGCTGTGCGGCCAGGAAACGCTAGGCTCCCTACTACGCAATA 1098	
QY	424	CGACGCCAGCTCGAATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTGGGG 483	
Db	1099	CGACGCCAGCTCGAATTTGCTGTTGGGGCGGCTACTTTCGTTCCGCTATGACGTGGGG 1158	
QY	484	GACCTCTGCGGATCTGTTCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCTCGCCGGCAT 543	
Db	1159	GATCTCTCGGATCTGTTTCTCTCTCGTCGCGCCAGCTGTTCACTTCTCGCCCGCCGCGCAT 1218	
QY	544	GAGACGCTGAGGACTGCAATGTCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 603	
Db	1219	GAGACGCTACAGGACTGCAATTTGTTCAATCTATCCCGGCCACATAACGGGTCAACCGCATG 1278	
QY	604	GCTTGGGATATGATGATGAACTGCTCGCTACAAACGGCCCTGGTGGTATCGCAGTGCTC 663	
Db	1279	GCTTGGGATATGATGATGAACTGCTCACCCACAGCAGCCCTAGTCTGATCGCAGTTACTC 1338	
QY	664	CGGATCCCAAGCTGTCGTGACATGTCGCGGGGGGCCCATTTGGGAGTCTCTGGCGGGT 723	
Db	1339	CGGATCCCAAGCTATCGTGATATGTCGCGGGGGGCCCATCTGGGAGTCTCTGGCGGGC 1398	
QY	724	CTCGCTACTATTCATGTCGGGAACTGGGCTAAGTTTGTATTTGATGATGATCTACTCTTT 783	
Db	1399	CTCGCTACTATTCATGTCGGGAACTGGGCTAAGTTTGTATTTGATGATGATCTACTCTTT 1458	
QY	784	GC 785	
Db	1459	GC 1460	
RESULT 15			
HCV132996 9418 bp RNA linear VRL 04-MAY-1999			
LOCUS Hepatitis C virus, complete genome, isolate HCV-AD78.			
DEFINITION HCV132996			
ACCESSION AJ132996			
VERSION AJ132996.1 GI:4753718			
KEYWORDS complete genome; polyprotein.			
SOURCE Hepatitis C virus			
ORGANISM Hepatitis C virus			
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
1. Rispeter, K.			
AUTHORS Thesis (1998) Universitaetsklinikum Essen, Institut fuer Virologie			
JOURNAL 2 (bases 1 to 9418)			
REFERENCE Rispeter, K.			
AUTHORS Direct Submission			
TITLE Submitted (07-MAR-1999) Rispeter K., Universitaetsklinikum Essen,			
JOURNAL Institut fuer Virologie, Hufelandstrasse 55, 45122 Essen, GERMANY			
FEATURES			
source			
1..9418 /organism="Hepatitis C virus"			
/vrlon			
/mol_type="genomic RNA"			
/isolate="HCV-AD78"			
/db_xref="taxon:11103"			
/country="Germany"			
/note="genotype 1b-without entire 3'NCR"			
341..9379 /codon_start=1			
/product="polyprotein"			
CDS			


```
QY 184 TTGGCTTTGCTGCTGCTGCTGACCGGTTCCAGCTTCCGCTTATGAAGTGGCAACGTTGTC 243
Db 875 TTGGCTTTGCTGCTGCTGCTGACCGATCCCAAGCTTCCGCTTATGAAGTGGCAACGTTGTC 934
QY 244 GGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGCGGAC 303
Db 935 GGGGTGTACCATGTACAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGACAGCGGAC 994
QY 304 ATGATCATGACACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCCGCTGC 363
Db 995 ATGATCATGACACACCCCGGGTGGTCCCTGCGTTCCGGAGGATAACTCCTCTCGCTGC 1054
QY 364 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGACGCCAGCGTCCCCACACGACAATA 423
Db 1055 TGGGTAGCGCTCACCCCGACGCTCGCGGCTAGGAAATGGCAGTGTCCCCACACGCGGATA 1114
QY 424 CGACGCCACGCTCGAATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTGGGG 483
Db 1115 CGACGCCACGCTCGAATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTGGGG 1174
QY 484 GACCTCTGGCGATGCTGCTTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCCGGCAT 543
Db 1175 GATTTTGGCGATGCTGTTTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCCGGCAT 1234
QY 544 GAGACGCTGACGAGCTGCAATTTGCTCAATCTATCCCGGCCACACATAACGGGTACCGTATG 603
Db 1235 GAGACAGTACAGGAATGTAAATGCTCAATTTACCCCGGCCACGTCACAGGTACCCGATG 1294
QY 604 GCTTGGGATATGATGATGAATGCTGCTGCTTACAACTGCTGCTGCTGCTGCTGCTGCTGCT 663
Db 1295 GCTTGGGATATGATGATGAATGCTGCTTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1354
QY 664 CGATCCCAACAAGCTGCTGAGACATGCTGGCGGGGCCCAATTGGGGAGTCTCTGGCGGGT 723
Db 1355 CGGATCCCAACAAGCTGCTGAGACATGCTGGCGGGGCCCAATTGGGGAGTCTCTGGCGGGC 1414
QY 724 CTGCGCTACTATTCCATGGTGGGAACTGGGCTTAAGGTTTTGATTGTGATGCTACTCTTT 783
Db 1415 CTGCGCTACTATTTCGATGGTGGGAACTGGGCTTAAGGTTTTGATTGTGATGCTACTCTTT 1474
QY 784 GCT 786
Db 1475 GCT 1477
```

Search completed: December 20, 2003, 02:01:50
Job time : 3012.09 secs

THIS PAGE BLANK (USPTO)

XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

proteins - in presence of di:sulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

Claim 23; Fig 21; 146pp; English.

AA12704-T12709 and AA12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence of the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

Query Match 100.0%; Score 795; DB 17; Length 795;
Best Local Similarity 100.0%; Pred. No. 3.6e-204;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATT 60
1 ATGTTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATT 60

61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCGGGGT 120
61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCGGGGT 120

121 CTGGAGGACGGGTCAACTATGCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTC 180
121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTC 180

181 CTCCTTGGCTTTGCTGTCTGTCTGACCGGTTCCAGCTTCGGCTTATGAAGTGGCAACGTG 240
181 CTCCTTGGCTTTGCTGTCTGTCTGACCGGTTCCAGCTTCGGCTTATGAAGTGGCAACGTG 240

241 TCGGGATGTACATGTGACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGCAGCG 300
241 TCGGGATGTACATGTGACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGCAGCG 300

301 GACATGATCATGCACACCCCGGGTCGCTGCCCTGCGTTTCGGGAGAACAACTCTTCCCGC 360
301 GACATGATCATGCACACCCCGGGTCGCTGCCCTGCGTTTCGGGAGAACAACTCTTCCCGC 360

361 TGCTGGGTAGCGTCAACCCCGACGCTGAGAACGCCAGCGTCCCCACACGACA 420
361 TGCTGGGTAGCGTCAACCCCGACGCTGAGAACGCCAGCGTCCCCACACGACA 420

421 ATACGACGCCACGTTCGATTTGCTGTGTGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
421 ATACGACGCCACGTTCGATTTGCTGTGTGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480

481 GGGGACCTCTCGGGAATCTGTTCTTCGTGTCGCCAGCTGTTACCATCTCGCTCCCGG 540
481 GGGGACCTCTCGGGAATCTGTTCTTCGTGTCGCCAGCTGTTACCATCTCGCTCCCGG 540

541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT 600
541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT 600

601 ATGGCTTGGGATATGATGATGAACGTGTCGCCCTTACAAACGGCCCTTGGTATCGCAGCTG 660


```
Db 1 ATGTGGGTAAGGTCAATCGATACCTTACATGCGGCTTCGCGGACCTCGTGGGTACATT 60
Qy 61 CCCTCGTTCGGCGCCCCCTAGGGGGCGCTGCCAGGCCCCCTGGCGCATGGCGTCCGGGTT 120
Db 61 CCCTCGTTCGGCGCCCCCTAGGGGGCGCTGCCAGGCCCCCTGGCGCATGGCGTCCGGGTT 120
Qy 121 CTGGAGACGGGTGAATGATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTC 180
Db 121 CTGGAGACGGGTGAATGATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTC 180
Qy 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGCACAGTG 240
Db 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGCACAGTG 240
Qy 241 TCCGGGATGATACCATGTCACGAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGC 300
Db 241 TCCGGGATGATACCATGTCACGAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGC 300
Qy 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCGCTGCGTTCGGGAGAACACTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCGCTGCGTTCGGGAGAACACTTCCCGC 360
Qy 361 TGCTGGGTAGCGCTCACCCACGCTCGCAGCTAGGAACGCGAGCGTCCCGACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCACGCTCGCAGCTAGGAACGCGAGCGTCCCGACACGACA 420
Qy 421 ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
Db 421 ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
Qy 481 GGGGACCTCTGGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GGGGACCTCTGGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CATGAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGGCCACATAAGGGTCACCGT 600
Db 541 CATGAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGGCCACATAAGGGTCACCGT 600
Qy 601 ATGGCTTGGGATATGATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 ATGGCTTGGGATATGATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 CTCGGATCCCAAGCTGCTGGAACATGCTGCGGGGGCCCAATGCGGAGTCTTGCGG 720
Db 661 CTCGGATCCCAAGCTGCTGGAACATGCTGCGGGGGCCCAATGCGGAGTCTTGCGG 720
Qy 721 GGTCTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGATGATGCTACTC 780
Db 721 GGTCTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGATGATGCTACTC 780
Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795
```

RESULT 3

AAL48939

ID AAL48939 standard; DNA; 2082 BP.

AC AAL48939;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

18-JUL-2002.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

11-JAN-2002; 2002WO-EF00219.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

New therapeutic vaccine compositions comprising at least one purified

recombinant hepatitis C virus (HCV) single or specific oligomeric

protein E1 or E2, useful for immunizing humans

from HCV infection

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 98.0%; Score 778.8; DB 24; Length 2082;

Best Local Similarity 99.7%; Pred. No. 1.1e-199;

Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

Db
484 GACCTCTGGGATCTGTCTTCCTCGTCTCCAGCTGTTACCATCTCGCTCGCGCGCAT 543
544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTACCGGTATG 603
QY
544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTACCGGTATG 603

db
544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCACCGGATG 600

604 GCATTGGGATATGATGATGAACCTGGTGCCTACAACGGGCCCTGGTGGTATCGCAGCTGCTC 663

db

604 GCCTGGGATATGATGATGAACCTGGTCGCCCTACAACGGCCCTGGTGGTATCGCAGTGCTC 663

654 CCGATCCCAAGCTGTCTGTGGACATGGTGGCGGGGCCCATTTGGGAGTCTCTGGCGGGT 723

[illegible][illegible]

724 CTCGCCCTACTATTCCATGGTGGGGAACCTGGGCTAAGGTTTGTGATTTGGAGTGCACACGTTT

Db
724 CTCGCCCTACTATTCCATGGTGGGGAACTGGGGCTAAGGTTTTGGTTGTGATGCTACTCTTT 783

Qv 784 GC 785

784 GC 785

RESULT 4

AAT12973
ID AAT12973 standard; DNA; 2086 BP.

XX
XX
XX

AC
AA112913;
XX

DT 24-SEP-1996 (first entry)
XX

DE HCV E1 construct HCCI65.

XX
KW

HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

KW serotype; reversed phase hybridisation assay; genotyping; energy; ss.
KW ss.

XX
Cc
XX
Hematitis C virus.

XX nepaculus 3 virens:

PN W09604385-A2.
XX

PD 15-FEB-1996.
yy

31-JUL-1995; 95WO-EP03031.

XX
PR 29-JUL-1994; 94EP-0870132.

XX
PA (INNO-) INNOGENETICS NV.

XX	De Martynoff G.	Maertens G.
XX	De Martynoff G.	Maertens G.
XX	De Martynoff G.	Maertens G.

PI
HOSMAN F, Buysse M, de Marchinova S, ...
XX

DR WPI; 1996-129401/13.
XX

Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

PT produce proteins suitable for direct use in vaccines or diagnostic

PT assays of HCV
XX

PS Claim 23; Fig 21; 146pp; English.

CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) ER
AA
AA These sequences are included in the sequence constructs

CC and E2 protein coding sequence constructs. These constructs
CC and E2 protein coding sequence constructs. These constructs
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins

The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by

carrying out a disulphide bond cleavage, or a reduction step with a

CC The constructs containing the purified HCV envelope proteins can be used
CC disulphide bond cleavage agents, and the detection of HCV
CC The constructs containing the purified HCV envelope proteins can be used
CC The constructs containing the purified HCV envelope proteins can be used

for vaccinating humans against HCV, for *in vitro* detection of new antibodies in a sample, and in a serotyping assay for detecting one or

more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated

[illegible]

XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection
XX
XX Example 2; Page 212-215; 243pp; English.
PS
XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunizing humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 other;
SQ
Query Match 96.6%; Score 767.8; DB 24; Length 2434;
Best Local Similarity 99.6%; Pred. No. 1.1e-196;
Matches 780; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 4 TTGGGTAGGTCATGATACCTTACATGCGCTTCGCGGACCTCGTGGGTACATTCGG 63
DB 355 TTGGGTAGGTCATGATACCTTACATGCGCTTCGCGGACCTCGTGGGTACATTCGG 414
QY 64 CTCGTCGGGCGCCCGCTAGGGGGCGCTGCCAGGGCGCTTGGGCGATGGCGTCCGGTCTG 123
DB 415 CTCGTCGGGCGCCCGCTAGGGGGCGCTGCCAGGGCGCTTGGGCGATGGCGTCCGGTCTG 474
QY 124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTCTC 183
DB 475 GAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTCTC 534
QY 184 TTGGCTTTGCTGTCCTG-TCTGACCGCTTCCAGCTTCCGCTTATGAAGTGGCGAACGTGTC 242
DB 535 TTGGCTTTGCTGTCCTG-TCTGACCGCTTCCAGCTTCCGCTTATGAAGTGGCGAACGTGTC 594
QY 243 CGGGATGTACCATGTCCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGA 302
DB 595 CGGGATGTACCATGTCCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGA 654
QY 303 CATGATCATGCACACCCCGGGTGGTCCGCTTCCGAGAGAACAACTCTTCCCGCTG 362
DB 655 CATGATCATGCACACCCCGGGTGGTCCGCTTCCGAGAGAACAACTCTTCCCGCTG 714
QY 363 CTGGGTAGCGCTCACCCCGGCTGCGAGTGGAAAGCGCGGTCCCAACCAACCAAT 422
DB 715 CTGGGTAGCGCTCACCCCGGCTGCGAGTGGAAAGCGCGGTCCCAACCAACCAAT 774
QY 423 ACGACGCGCATGCTGCTTGGGGGGCTGCTTTCTGTTCCGCTATGATGCTGGG 482
DB 775 ACGACGCGCATGCTGCTTGGGGGGCTGCTTTCTGTTCCGCTATGATGCTGGG 834
QY 483 GGACCTCTGCGGATCTGCTTCTCTGCTCCAGCTGTTTCCATCTGCGCTCGCGGCA 542
DB 835 GGACCTCTGCGGATCTGCTTCTCTGCTCCAGCTGTTTCCATCTGCGCTCGCGGCA 894
QY 543 TGAGACGGTGCAGGCTGCAATCTCAATCTATCCCGGCACATAACGGGTCAACGGTAT 602
DB 895 TGAGACGGTGCAGGCTGCAATCTCAATCTATCCCGGCACATAACGGGTCAACGGTAT 954
QY 603 GCCTTGGGATATGATGAACTGGTGGCTTACAAACGGCCCTGGTGTATGCGAGTGTCT 662
DB 955 GCCTTGGGATATGATGAACTGGTGGCTTACAAACGGCCCTGGTGTATGCGAGTGTCT 1014
QY 663 CCGGATCCCAAGCTGCTGGGACATGCTGGCGGGGCCATTTGGGGAGTCTCTGGCGGG 722

Db 1015 CCGATCCCAAGCTGCTGTCGACATGTTGGCGGGGCCCAATGGGGAGTCTCTGGCGG 1074
QY 723 TCTGCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGGATTGTGATGCTACTCT 782
Db 1075 CTTGCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGGATTGTGATGCTACTCT 1134
QY 783 TGC 785
Db 1135 TGC 1137
RESULT 7
ASK91411
ID ABK91411 standard; DNA; 9605 BP.
XX
AC ABK91411;
XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..9374
FT /*tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT
XX
PN WO200259321-A2.
XX
PD 01-AUG-2002.
XX
PP 16-JAN-2002; 2002WO-BP00526.
XX
PR 23-JAN-2001; 2001US-263479P.
XX
PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
PI De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI: 2002-599793/64.
DR P-PSDB; ABG32451.
XX
DR New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression -
XX
PS Claim 9; Page 36-39; 69pp; English.
XX
CC The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV

CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is the HCV replicon Con 1, used as a basis for
CC the adaptive mutations of the invention.

XX
SQ Sequence: 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 other;
Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGTAAGGTTCATCGATACCTTACATCGCGCTTCGCGACCTCGGGGTACATCCG 63
DB |||||
QY 64 CTGTCGCGCCCCCTAGGCGCGCTGCCAGGCGCTTGGCGCATGCGCGGTCTCTG 123
DB |||||
QY 756 CTGTCGCGCCCCCTAGGCGCGCTGCCAGGCGCTTGGCGCATGCGCGGTCTCTG 815
DB |||||
QY 124 GAGACGCGGTGAATCATATGCAAGGGAATTTGCGCGGTCTCTTCTATCTCTC 183
DB |||||
QY 816 GAGACGCGGTGAATCATATGCAAGGGAATTTGCGCGGTCTCTTCTATCTCTC 875
DB |||||
QY 184 TTGCTTTGCTGCTCTCTCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGCATCTCC 243
DB |||||
QY 876 TTGCTTTGCTGCTCTCTCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGCATCTCC 935
DB |||||
QY 244 GGGATGTACCATGTTCACGACGACTGCTTCAAGCATTTGATGAGGACGCGGAC 303
DB |||||
QY 936 GGAGTGTACCATGTTCACGACGACTGCTTCAAGCATTTGATGAGGACGCGGAC 995
DB |||||
QY 304 ATGATCATGACACCCCGGCTGCGCTTCCAGCTTCGCGGTCTCTTCTCTCTCTC 363
DB |||||
QY 996 ATGATCATGACACCCCGGCTGCGCTTCCAGCTTCGCGGTCTCTTCTCTCTCTC 1055
DB |||||
QY 364 TGGTAGCGTTCACCCCGGCTGCGCTTCCAGCTTCGCGGTCTCTTCTCTCTCTC 423
DB |||||
QY 1056 TGGTAGCGTTCACCCCGGCTGCGCTTCCAGCTTCGCGGTCTCTTCTCTCTCTC 1115
DB |||||
QY 424 CGACGCCACGTCGATTTGCTGCTTGGCGCGCTCTTCTGTTCCGCTATGTAGTGGG 483
DB |||||
QY 1116 CGACGCCACGTCGATTTGCTGCTTGGCGCGCTCTTCTGTTCCGCTATGTAGTGGG 1175
DB |||||
QY 484 GACCTCTGCGGATGCTCTCTCTGCTTCCAGCTTTCACCACTCGCGCTCGCGGAT 543
DB |||||
QY 1176 GATCTCTGCGGATGCTCTCTCTGCTTCCAGCTTTCACCACTCGCGCTCGCGGAT 1235
DB |||||
QY 544 GAGACGCTGACGACTGCAATTTGCTGCTTCAATTTGCTGCTTCAATTTGCTGCT 603
DB |||||
QY 1236 GAGACGCTGACGACTGCAATTTGCTGCTTCAATTTGCTGCTTCAATTTGCTGCT 1295
DB |||||
QY 604 GCTTGGGATATGATGAACTGCTGCTTCAATTTGCTGCTTCAATTTGCTGCTTCA 663
DB |||||
QY 1296 GCTTGGGATATGATGAACTGCTGCTTCAATTTGCTGCTTCAATTTGCTGCTTCA 1355
DB |||||
QY 664 CGGATCCACAAAGCTGCTGAGCATGCTGCGCGGCGCCATTTGGGAGTCTCTGCGG 723
DB |||||
QY 1356 CGGATCCACAAAGCTGCTGAGCATGCTGCGCGGCGCCATTTGGGAGTCTCTGCGG 1415
DB |||||
QY 724 CTGCTCTACTATTCATGCTGGGAACTGGGCTTGAAGGTTTGTATGATGCTACTCTT 783
DB |||||
QY 1416 CTGCTCTACTATTCATGCTGGGAACTGGGCTTGAAGGTTTGTATGATGCTACTCTT 1475
DB |||||
QY 784 GC 785
DB ||
QY 1476 GC 1477
DB ||

RESULT 8
ABK91424
ID ABK91424 standard; DNA; 9605 BP.
XX
AC ABK91424;

XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA mutant 1.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
OS Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
PH CDS 342..9374
FT /*tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT mutation replace (3625,G)
FT /*tag= b
XX
XX W0200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-EP00526.
XX
XX 23-JAN-2001; 2001US-263479P.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS3 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression
XX
XX Claim 9; Page -: 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is an HCV replicon Con 1 mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the HCV sequence appearing as
CC ABK91411 and the information in Claim 9.
XX
SQ Sequence 9605 BP; 1910 A; 2883 C; 2732 G; 2079 T; 0 other;

Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGTAAGGTTCATCGATACCTTACATCGCGCTTCGCGACCTCGGGGTACATCCG 63

Mon Dec 22 13:28:52 2003

Db	696	TTGGGTAAGGTATCATGATACCTTACGTCGGCTTCGCCGATCTCATGGGTACATTCGG	755
Qy	64	CTGTCGGCGCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	123
Db	756	CTGTCGGCGCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	815
Qy	124	GAGGACGGCGTGAACATATGACAAAGGGAATTTCCCGGTTGCTTCTCTATCTTCCTC	183
Db	816	GAGGACGGCGTGAACATATGACAAAGGGAATTTCCCGGTTGCTTCTCTATCTTCCTC	875
Qy	184	TTGCTTTGCTGCTCTCTGACCGTTCAGCTTATGAGTTCGGCTTATGAGTTCGGCTTC	243
Db	876	TTGCTTTGCTGCTCTCTGACCGTTCAGCTTATGAGTTCGGCTTATGAGTTCGGCTTC	935
Qy	244	GAGGATGATACCATGTCACGAAAGCATGCTCCAACTCAAGCATTTGTGTATGAGGACGGAC	303
Db	936	GAGGATGATACCATGTCACGAAAGCATGCTCCAACTCAAGCATTTGTGTATGAGGACGGAC	995
Qy	304	ATGATCATGACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGCTGC	363
Db	996	ATGATCATGACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGCTGC	1055
Qy	364	TGGGTAGCGCTCACCCACACGCTCGACAGTAAAGGAGCGAGCGTCCCGACGACATA	423
Db	1056	TGGGTAGCGCTCACCCACACGCTCGACAGTAAAGGAGCGAGCGTCCCGACGACATA	1115
Qy	424	CGACGCCAGCTGATTTGCTGTTGGGGGCGCTCTTCTGCTCCGCTATGATGATGGG	483
Db	1116	CGACGCCAGCTGATTTGCTGTTGGGGGCGCTCTTCTGCTCCGCTATGATGATGGG	1175
Qy	484	GACCTCTGGGATGCTGCTTCTCTGCTCCAGCTGTTTCCACATCTGCGCTCGCGGCAT	543
Db	1176	GATCTCTGGGATGCTGCTTCTCTGCTCCAGCTGTTTCCACATCTGCGCTCGCGGCAT	1235
Qy	544	GAGACGGTCAGAGCTGCAATTTGCTCAATATCCCGGCCACATAAGCGGTACCGTATG	603
Db	1236	GAGACAGTACAGAGCTGCAATTTGCTCAATATCCCGGCCACATGACAGGTACCGTATG	1295
Qy	604	GCTTGGGATGATGATGAATGCTGCTCCCTACAGCGCCCTGGTGGTATCGAGTGTCTC	663
Db	1296	GCTTGGGATGATGATGAATGCTGCTCCCTACAGCGCCCTGGTGGTATCGAGTGTCTC	1355
Qy	664	CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCGCCATTTGGGAGTCTGCGGGT	723
Db	1356	CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCGCCATTTGGGAGTCTGCGGGT	1415
Qy	724	CTGCGCTACTATTCCATGTTGGGAACTGCGGCTTAAGGTTTGTATGATGATCTCTTT	783
Db	1416	CTTGCCTACTATTCCATGTTGGGAACTGCGGCTTAAGGTTTGTATGATGATCTCTTT	1475
Qy	784	GC 785	
Db	1476	GC 1477	
RESULT 9	ABK91425	standard; DNA; 9605 BP.	
ID	ABK91425		
XX	AC	ABK91425;	
XX	DT	15-NOV-2002 (first entry)	
XX	DE	Hepatitis C virus Con 1 isolate DNA mutant 2.	
XX	KW	HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;	
XX	KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;	
XX	KW	internal ribosome entry site; IRES; NS5A; HCV replication; mutant.	
XX	OS	Hepatitis C virus.	
XX	OS	Synthetic.	

H	Key	Location/Qualifiers
CDS		342..9374
T		/*tag= a
T		/product= "HCV polyprotein"
T		/note= "The polyprotein consists of the Core, E1,
T		E2, P7, NS2, NS3, NS4B, NS4A, NS5A and NS5B proteins"
T		replace (3946,A)
T	mutation	/*tag= b
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		
T		

Db 876 TTGGCTTTGCTGCTGCTTTGACATCCAGCTTCCGCTTATGAAGTGCACAGTATCC 935
 Qy 244 GGATGTACCATGTCAGAACGAGTGTCTCAACTCAAGCATTTGTATGAGGACGGAC 303
 Db 936 GGATGTACCATGTCAGAACGAGTGTCTCAACTCAAGCATTTGTATGAGGACGGAC 995
 Qy 304 ATGATCATGACACCCCGGGTGGTCCCTGCTGCTTGGGAGAACAACTCTTCCCGTGC 363
 Db 996 ATGATCATGACACCCCGGGTGGTCCCTGCTGCTTGGGAGAACAACTCTTCCCGTGC 1055
 Qy 364 TGGGTAGCGTCAACCCCGAGCTCGACGCTAGGAAGCCAGCGTCCACACGACAAATA 423
 Db 1056 TGGGTAGCGTCAACCCCGAGCTCGACGCTAGGAAGCCAGCGTCCACACGACAAATA 1115
 Qy 424 CGACGCCATGCGATTTGCTGTTGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 483
 Db 1116 CGACGCCATGCGATTTGCTGTTGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGA 1175
 Qy 484 GACCTCTGGGATGCTGTTCTGCTGCTCCAGCTGTTCACCATCTCGCTCGCGGCAT 543
 Db 1176 GATCTCTGGCGATGCTGTTCTGCTGCTCCAGCTGTTCACCATCTCGCTCGCGGCAC 1235
 Qy 544 GAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 603
 Db 1236 GAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 1295
 Qy 604 GCTTGGGATATGATGAATGCTGCTCCCTACACCGCCCTGGTGTGTATCGCAGTCTCT 663
 Db 1296 GCTTGGGATATGATGAATGCTGCTCCCTACACCGCCCTGGTGTGTATCGCAGTCTCT 1355
 Qy 664 CGATCCCAAGCTGTGTGACATGTTGGCGGGGCCCATTTGGGAGTCTTGGCGGGT 723
 Db 1356 CGATCCCAAGCTGTGTGACATGTTGGCGGGGCCCATTTGGGAGTCTTGGCGGGC 1415
 Qy 724 CTGCGCTACTATCCATGTTGGGAGTGTGCTTGGGAGTCTTGGGAGTCTTGGGAGT 783
 Db 1416 CTGCGCTACTATCCATGTTGGGAGTGTGCTTGGGAGTCTTGGGAGTCTTGGGAGT 1475
 Qy 784 GC 785
 Db 1476 GC 1477

RESULT 10

ABK91426
 ID ABK91426 standard; DNA; 9605 BP.
 XX AC ABK91426;
 XX DT 15-NOV-2002 (first entry)
 XX DE Hepatitis C virus Con 1 isolate DNA mutant 3.
 XX KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
 XX OS Hepatitis C virus.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT CDS 342..9374
 FT /*tag= a
 FT /product= "HCV polyprotein"
 FT /note= "The polyprotein consists of the Core, E1,
 FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 FT mutation replace (4380, G)
 FT /*tag= b
 FT WO200259321-A2.
 XX PN 01-AUG-2002.
 XX PD

XX 16-JAN-2002; 2002MO-EP00526.
 XX 23-JAN-2001; 2001US-263479P.
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 XX De Francesco R, Migliaccio G, Paonessa G;
 XX WPI; 2002-599793/64.
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication
 PT and expression
 XX Claim 9; Page -: 69pp; English.
 XX The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the
 CC mutations are detailed in the specification. Also included are
 CC (1) an expression vector comprising a nucleotide sequence coding for
 CC the altered nucleic acids, which is transcriptionally coupled to an
 CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
 CC the altered nucleic acids; (3) a recombinant cell produced by introducing
 CC into a human hepatoma cell the altered nucleic acids; (4) producing a
 CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
 CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
 CC method; and (6) measuring the ability of a compound to affect HCV
 CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
 CC studying HCV replication and expression, and HCV and host cell
 CC interactions, producing HCV RNA and proteins, and providing a system
 CC for measuring the ability of a compound to modulate one or more HCV
 CC activities e.g. to discover drugs which may treat HCV mediated
 CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
 CC The present sequence is an HCV replicon Con 1 mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the HCV sequence appearing as
 CC ABK91411 and the information in Claim 9.
 XX Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 other;
 SQ Query Match 88.5%; Score 703.6; DB 24; Length 9605;
 Best Local Similarity 93.7%; Pred. No. 3.4e-179;
 Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 Qy 4 TTGGGTAAGTGTATCGATACCCCTTACATCGCGCTTCGCCGACCTCTGTTGGGTACATTCCG 63
 Db 696 TTGGGTAAGTGTATCGATACCCCTTACATCGCGCTTCGCCGACCTCTGTTGGGTACATTCCG 755
 Qy 64 CTGCTGGCGCCCGCCCTAGGGGGGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTCTG 123
 Db 756 CTGCTGGCGCCCGCCCTAGGGGGGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTCTG 815
 Qy 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTCTC 183
 Db 816 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTCTC 875
 Qy 184 TTGGCTTTGCTGCTCTGCTGTGACCGTTCCAGCTTTCGCTTATGAAGTGCACAGTGTCC 243
 Db 876 TTGGCTTTGCTGCTCTGCTGTGACCGTTCCAGCTTTCGCTTATGAAGTGCACAGTGTCC 935
 Qy 244 GGGATGTACCATGTACGAAACGATGCTCCAACTCAAGCATTTGTATGAGGACGGAC 303
 Db 936 GGGATGTACCATGTACGAAACGATGCTCCAACTCAAGCATTTGTATGAGGACGGAC 995
 Qy 304 ATGATCATGACACCCCGGGTGGTCCCTGCTGCTTGGGAGAACAACTCTTCCCGTGC 363
 Db 996 ATGATCATGACACCCCGGGTGGTCCCTGCTGCTTGGGAGAACAACTCTTCCCGTGC 1055
 Qy 364 TGGGTAGCGTCAACCCCGAGCTCGACGCTAGGAAGCCAGCGTCCACACGACAAATA 423

1056	db	TGGGTAGCCTCACTCCACGCTCGGGCCAGGAAGCTAGCGTCCCACTACGACGATA	1115
424	QY	CGACGCCACGTCGAATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTGGGG	483
1116	db	CGACGCCATGTCGAATTTGCTCGTTGGGGCGGCTGCTCTCTGCTCCGCTATGTACGTGGGA	1175
484	QY	GACCTCTGGGATCTGCTTTCTTCGTCCTCCAGCTGTTCCACCATCTCGCCTCGCGGCAT	543
1176	db	GATCTCTCGGATCTGTTTCTTCGTCGCCAGCTGTTCACTTCTTCGCTCGCGGCAC	1235
544	QY	GAGACGGTCGAGCTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACGTATG	603
1236	db	GAGACAGTACAGGACTGCAATTGCTCAATATATCCCGGCCACGTGACAGGTCAACCGTATG	1295
604	QY	GCTTGGGATATGATGAACTGGTCGCCCTACAAACGGCCCTGGTGGTATCGCAGCTGCTC	663
1296	db	GCTTGGGATATGATGAACTGGTCACCTACAGCAGCCCTAGTGGTATCGCAGTTACTC	1355
664	QY	CGGATCCCAACAGCTGTCGTGGACATGGTGGCGGGGCCCATTTGGGAGTCCTGGCGGGT	723
1356	db	CGGATCCCAACAGCTGTCGTGGATATGGTGGCGGGGCCCATTTGGGAGTCTTAGCGGGC	1415
724	QY	CTCGCCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGTATGTGATGCTACTCTTT	783
1416	db	CTTGCCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGTATGTGATGCTACTCTTT	1475
784	QY	GC 785	
		1476 GC 1477	
	db		

XX	RESULT 11	
XX	ABK91428	
XX	ID	ABK91428 standard; DNA; 9605 BP.
XX	AC	ABK91428;
XX	DT	15-NOV-2002 (first entry)
XX	DE	Hepatitis C virus Con 1 isolate DNA mutant 5.
XX	KW	HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
XX	KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX	KW	internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
XX	OS	Hepatitis C virus.
XX	OS	Synthetic.
XX	PH	
XX	Key	Location/Qualifiers
FT	CDS	342..9374
FT		/*tag= a
FT		/product= "HCV polypotein"
FT		/note= "The polypotein consists of the Core, E1,
FT		E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT	mutation	replace (6463,A)
FT		/*tag= b
XX		
XX	W0200259321-A2.	
XX		
XX	01-AUG-2002.	
XX		
XX	16-JAN-2002; 2002WO-EP00526.	
XX		
XX	23-JAN-2001; 2001US-263479P.	
XX		
XX	(RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.	
XX		
XX	De Francesco R, Migliaccio G, Paonessa G;	
XX		
XX	WPI; 2002-599793/64.	
XX		
XX	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV	

NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region, useful in studying HCV replication and expression -

Claim 9; Page -; 69pp; English.

The invention relates to nucleic acid molecules comprising altered HCV NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region coding for one or more NS3, NS4A, or EMCV IRES mutations, respectively. The location of the mutations are detailed in the specification. Also included are (1) an expression vector comprising a nucleotide sequence coding for the altered nucleic acids, which is transcriptionally coupled to an exogenous promoter; (2) a recombinant cell human hepatoma cell comprising the altered nucleic acids; (3) a recombinant cell produced by introducing into a human hepatoma cell the altered nucleic acids; (4) producing a HCV (hepatitis C virus) replicon enhanced cell or which containing a functional HCV replicon; (5) an HCV replicon enhanced cells made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicons and HCV replicon enhanced cells are useful in studying HCV replication and expression, and HCV and host cell interactions, producing HCV RNA and proteins, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated diseases such as liver failure, cirrhosis and hepatocellular carcinoma. The present sequence is an HCV replicon Con 1 mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the HCV sequence appearing as follows and the information in Claim 9.

XX	Sequence	9605 BP	1909 A	2884 C	2733 G	2079 T	0 other
XX	Query Match	88.5%	Score 703.6	DB 24	Length 9605		
XX	Best Local Similarity	93.7%	Pred. No. 3.4e-179				
XX	Matches 733	Conservative	0	Mismatches 49	Indels 0	Gaps 0	
QY	4	TTGGGTAAAGTCATCGATACCCCTTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCGG	63				
DB		TTGGGTAAAGTCATCGATACCCCTCACGTGCGGCTTCGCCGATCTCATGGGTACATTCGG	755				
QY	64	CTCGTGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTTG	123				
DB		CTCGTGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTTG	815				
QY	124	GAGGACGGGTGAACATATGCAACAGGGAAATTGCGCGGTGCTCTTCTCTACTTCCTC	183				
DB		GAGGACGGGTGAACATATGCAACAGGGAAATCTGCCCGTTGCTCTCTTTCTATCTTCCTT	875				
QY	184	TTGGCTTTGTCGTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCAGCAAGTGTCC	243				
DB		TTGGCTTTGTCGTCTGCTTGAACCATCCAGTTCGCGTTATGAAGTGCAGCAAGTGTATCC	935				
QY	244	GGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCATGTGTATGAGGAGCGGAC	303				
DB		GGGATGTACCATGTACAGAACGACTGCTCCAAACGCAAGCATGTGTATGAGGAGCGGAC	995				
QY	304	ATGATCATGACACCCCGGGTGCCTGCGTTCCGGAGAACAACTCTTCCTCCGCTGC	363				
DB		ATGATCATGATACCCCGGGTGCCTGCGTTCCGGAGAACAACTCTCTCCGCTGC	1055				
QY	364	TGGGTAGCGCTACCCCCACGCTCGCAGCTAGGAAACGCGTCCCAACACGACATA	423				
DB		TGGGTAGCGCTACCTCCACGCTCGCGGCGAGAACGCTAGCGTCCCAACGACGATA	1115				
QY	424	CGACGCCACGTCGATTTGCTCGTTGGGGGGCTGCTTCTGTTTCGCTATGTACGTGGGG	483				
DB		CGACGCCATGTCGATTTGCTCGTTGGGGGGCGCTGCTCTCTGCTCGCTATGTACGTGGGA	1175				
QY	484	GACCTCTCGGATCTGCTTCTCCTCGTCTCCAGCTGTTCAACATCTCGCTCCCGGCA	543				
DB		GATCTCTCGGATCTGTTTCTCCTCGTCCGCCAGCTGTTTCACTCTCTCGCTCCCGGCA	1235				
QY	544	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACG	603				

Db 1236 GAGACAGTACAGGACTGCAATTTGCTCAATATATCCCGGCCACGTGACAGTCCACGGTATG 1295
 QY 604 GCTTGGGATATGATGAACTGGTGCCTTACAAACGGCCCTGGTGGTATCGCAGCTGCTC 663
 Db 1296 GCTTGGGATATGATGAACTGGTGCCTTACAAACGGCCCTGGTGGTATCGCAGTATC 1355
 QY 664 CGATCCCAACAGTCTGCTGGACATGGTGGGGGGCCCATTTGGGGAGTCTGGCGGGT 723
 Db 1356 CGATCCCAACAGTCTGCTGGATATGGTGGGGGGCCCATTTGGGGAGTCTGGCGGGC 1415
 QY 724 CTGCTTACTATTCCATGGTGGGAACTGGGCTAAAGGTTTGGATGTGATCTCTTT 783
 Db 1416 CTGCTTACTATTCCATGGTGGGAACTGGGCTAAAGGTTTGGATGTGATCTCTTT 1475
 QY 784 GC 785
 Db 1476 GC 1477

RESULT 12

ABK91429
 ID ABK91429 standard; DNA; 9605 BP.
 XX
 AC ABK91429;
 DT 15-NOV-2002 (first entry)
 XX
 DE Hepatitis C virus Con 1 isolate DNA mutant 6.
 XX
 KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
 XX
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 FH Key
 CDS 342..9374
 FT Location/Qualifiers
 FT /tag= a
 FT /product= "HCV polyprotein"
 FT /note= "The polyprotein consists of the Core, E1,
 FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 FT mutation
 FT replace (6859,C)
 FT /tag= b
 XX
 PN WO200259321-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 16-JAN-2002; 2002WO-EP00526.
 XX
 PR 23-JAN-2001; 2001US-263479P.
 XX
 PA (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.
 XX
 PI De Francesco R, Migliaccio G, Paonessa G;
 XX WPI; 2002-599793/64.
 XX
 DR New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication
 PT and expression
 XX
 PS Claim 9; Page -: 69pp; English.
 XX
 CC The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the
 CC mutations are detailed in the specification. Also included are
 CC (1) an expression vector comprising a nucleotide sequence coding for

the altered nucleic acids, which is transcriptionally coupled to an
 exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
 the altered nucleic acids; (3) a recombinant cell produced by introducing
 into a human hepatoma cell the altered nucleic acids; (4) producing an
 HCV (hepatitis C virus) replicon enhanced cell or which containing a
 functional HCV replicon; (5) an HCV replicon enhanced cells made in the
 method; and (6) measuring the ability of a compound to affect HCV
 activity. The HCV replicons and HCV replicon enhanced cells are useful in
 studying HCV replication and expression, and HCV and host cell
 interactions, producing HCV RNA and proteins, and providing a system
 for measuring the ability of a compound to modulate one or more HCV
 activities e.g. to discover drugs which may treat HCV mediated
 diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
 Note: The present sequence is an HCV replicon Con 1 mutant of the invention.
 Note: The present sequence is not shown in the specification but
 was created by the indexer using the HCV sequence appearing as
 CC ABK91411 and the information in Claim 9.
 XX
 SQ Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 other;

Query Match 88.5%; Score 703.6; DB 24; Length 9605;
 Best Local Similarity 93.7%; Pred. No. 3.4e-179;
 Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTCAATCGATACCTTACATGCGGCTTCCGCGACCTCGTGGGTACATCCG 63
 Db 696 TTGGGTAAAGTCAATCGATACCTTACATGCGGCTTCCGCGACCTCGTGGGTACATCCG 755
 QY 64 CTGCTGGCGGCCCTCTAGGGGGCGCTCCAGAGGCTGCGCGATGCGCGTTCG 123
 Db 756 CTGCTGGCGGCCCTCTAGGGGGCGCTCCAGAGGCTGCGCGATGCGCGTTCG 815
 QY 124 GAGGACGGGTGAACATATGCAACAGGGAATTTCCCGGGTTCCTTTCTATCTTCCTC 183
 Db 816 GAGGACGGGTGAACATATGCAACAGGGAATTTCCCGGGTTCCTTTCTATCTTCCTC 875
 QY 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGCTTATGAGTGGCGCAACGTGTC 243
 Db 876 TTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGCTTATGAGTGGCGCAACGTGTC 935
 QY 244 GGGATGTACCATCTCAGACGACTGCTCCAACTCAAGCAATTTGATAGGAGCGCGAC 303
 Db 936 GGGATGTACCATCTCAGACGACTGCTCCAACTCAAGCAATTTGATAGGAGCGCGAC 995
 QY 304 ATGATCATGACACACCCCGGGTGGCTGCGCTTCCGAGAGCAACTCTCCCGCTGC 363
 Db 996 ATGATCATGACATACCCCGGGTGGCTGCGCTTCCGAGAGCAACTCTCCCGCTGC 1055
 QY 364 TGGGTAGCGCTCACCCCGCGCTGCGAGTGAAGAACGCCAGCGTCCCAACAGACAATA 423
 Db 1056 TGGGTAGCGCTCACCCCGCGCTGCGAGTGAAGAACGCCAGCGTCCCAACAGACAATA 1115
 QY 424 CGACGCCAGCTCGATTTGCTGCTGGGGGGCTGCTTCTGTTCCGCTATGTAGTGGGG 483
 Db 1116 CGACGCCAGCTCGATTTGCTGCTGGGGGGCTGCTTCTGTTCCGCTATGTAGTGGGG 1175
 QY 484 GACCTCTGCGGATCTGCTTCTGCTCCAGCTGTTCCACCATCTCGCTCGCGCGCAT 543
 Db 1176 GATCTCTGCGGATCTGCTTCTGCTCCAGCTGTTCCACCATCTCGCTCGCGCGCAT 1235
 QY 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATG 603
 Db 1236 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATG 1295
 QY 604 GCTTGGGATATGATGAACTGGTGGGAGTGGTGGGGGGCCCATTTGGGGAGTCTGCTC 663
 Db 1296 GCTTGGGATATGATGAACTGGTGGGAGTGGTGGGGGGCCCATTTGGGGAGTCTGCTC 1355
 QY 664 CGATCCCAACAGTCTGCTGGACATGGTGGGGGGCCCATTTGGGGAGTCTGCTC 723
 Db 1356 CGATCCCAACAGTCTGCTGGAGTATGGTGGGGGGCCCATTTGGGGAGTCTGCTC 1415
 QY 724 CTGCTTACTATTCCATGGTGGGAACTGGGCTAAAGGTTTGGATGTGATCTCTTT 783

CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma
CC The present sequence is an HCV replicon Con 1 mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the HCV sequence appearing as
CC was disclosed the information in Claim 9.

XX 2080 T; 0 other;

```
Query Match      88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%;
Pred. No. 3.4e-179;
Matches 733; Conservative 0; Mismatches 49; Indels 0;
```

TP ABK91430 standard; DNA; 9605 DF.

ABK91430
ID ABK91430 standard: DNA; 9605 BP.

AA
AC
ARK91430:

15-NOV-2002 (first entry)

XX
 isolate DNA mutant 7.

XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW HCV replication; mutant.

XX
OS
Hepatitis C virus.

03 Synthetic.

XX	Location/Qualifiers
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

EH	key	EH	CDS
	DOCACTO2/		342..9374

```

/*tag= a

```

```

FT      /product="HCV polyprotein"
FT      /note="The polyprotein consists of the Core, E1,
FT      E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT      replace (6931,C)
FT      /tag= b
FT      mutation

```

XX PN W0200259321-A2.

XX 01 2125-2002

XX
000000-EP00526

16-JAN-2002; 2002W0-EE00320.

23-JAN-2001; 2001US-263479P.

XX

NOT FOOT AND ANGEL FURT

PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELE

XX Do Francesco R. Migliaccio G, Paonessa G;
 07 Do Francesco R. Migliaccio G, Paonessa G;

XX
De Francisco R. Argente

DR WPI; 2002-599793/64.

XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression -
PT

XX
20
claim 8. page - : 69pp: English.

The invention relates to nucleic acid molecules comprising altered HCV NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region coding for one or more NS3, NS5A, or EMCV IRES mutations, respectively. The location of the mutations are detailed in the specification. Also included are (1) an expression vector comprising a nucleotide sequence coding for the altered nucleic acids, which is transcriptionally coupled to an exogenous promoter; (2) a recombinant cell human hepatoma cell comprising the altered nucleic acids; (3) a recombinant cell produced by introducing into a human hepatoma cell the altered nucleic acid; (4) producing an HCV (hepatitis C virus) replicon enhanced cell or which containing a functional HCV replicon; (5) an HCV replicon enhanced cells made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicons and HCV replicon enhanced cells are useful in studying HCV replication and expression, and HCV and host cell interactions, producing HCV RNA and proteins, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated

AC	ABK91431;	QY	4	TTGGGTAAAGTCAATGATACCCCTTACATCGCGGCTTCGCCGACCTCGTGGGGTACATTCGG	63
XX		DB	696	TTGGGTAAAGTCAATGATACCCCTTACATCGCGGCTTCGCCGACCTCGTGGGGTACATTCGG	755
DE	15-NOV-2002 (first entry)	QY	64	CTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGCGCATGCGCTCGGGGTTCTG	123
KW	Hepatitis C virus Con 1 isolate DNA mutant 8.	DB	756	CTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGCGCATGCGCTCGGGGTTCTG	815
KW	HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;	QY	124	GAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTGCTCTTTCTCTATCTTCCTC	183
KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;	DB	816	GAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTGCTCTTTCTCTATCTTCCTC	875
KW	internal ribosome entry site; IRES; NS5A; HCV replication; mutant.	QY	184	TTGGCTTTGCTGCTGCTGCTGACCCGTTCCAGCTTCGCTTATGAAGTGGCAACGTTGTC	243
OS	Hepatitis C virus.	DB	876	TTGGCTTTGCTGCTGCTGCTGACCCGTTCCAGCTTCGCTTATGAAGTGGCAACGTTGTC	935
OS	Synthetic.	QY	244	GGGATGACCATGTACAGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCGAC	303
PH	Location/Qualifiers	DB	936	GGGATGACCATGTACAGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCGAC	995
FT	342..9374	QY	304	ATGATCATGCACACCCCGGGTGCCTGCTGCGGTTCGGGAGAACAACTCTTCCCGCTGC	363
FT	/*tag= a	DB	996	ATGATCATGCATACCCCGGGTGCCTGCTGCGGTTCGGGAGAACAACTCTTCCCGCTGC	1055
FT	/product= "HCV polyprotein"	QY	364	TGGGTAGCGCTCACCCCGACGCTCGAGCTAGGACGCCAGCTGCCACCAAGCAATA	423
FT	/notes= "The polyprotein consists of the Core, E1,	DB	1056	TGGGTAGCGCTCACCCCGACGCTCGAGCTAGGACGCCAGCTGCCACCAAGCAATA	1115
FT	E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"	QY	424	CGACGCCAGTGCATTTGCTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTTGGG	483
FT	replace (6934,T)	DB	1116	CGACGCCAGTGCATTTGCTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTTGGG	1175
FT	/*tag= . b	QY	484	GACCTCTGCGGATCTGTTCTCTCTCCAGCTGTTTCCACCATCTCGCTCCCGGCGAT	543
PN	WO200259321-A2.	DB	1176	GATCTCTGCGGATCTGTTCTCTCTCCAGCTGTTTCCACCATCTCGCTCCCGGCGAT	1235
PD	01-AUG-2002.	QY	544	GAGACGGTCGAGGACTGCAATTTGCTCAATCTATCCCGGCACATAACGGGTACCGTATG	603
PF	16-JAN-2001; 2002WO-BF00526.	DB	1236	GAGACAGTACAGGACTGCAATTTGCTCAATCTATCCCGGCACATAACGGGTACCGTATG	1295
PR	23-JAN-2001; 2001US-263479P.	QY	604	GCTTGGGATATGATGATGAAGTGTGCTGCTACAAACGGCCCTGGTGTATCGCAGCTGCTC	663
XX	(RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.	DB	1296	GCTTGGGATATGATGATGAAGTGTGCTGCTACAAACGGCCCTGGTGTATCGCAGCTGCTC	1355
XX	De Francesco R, Migliaccio G, Paonessa G;	QY	664	CGGATCCCAAGCTGCTGTCGACATGTCGGCGGGGCCCATTTGGGGAGTCTCTGGCGGGT	723
XX	WPI; 2002-599793/64.	DB	1356	CGGATCCCAAGCTGCTGTCGACATGTCGGCGGGGCCCATTTGGGGAGTCTCTGGCGGGT	1415
PT	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV	QY	724	CTCGCTACTATTCATGTTGGGAACTGGGCTAAGGTTTGTATGTTGATGCTACTCTTT	783
PT	NS5 encoding region, or encephalomyocarditis virus (EMCV) internal	DB	1416	CTTGCCTACTATTCATGTTGGGAACTGGGCTAAGGTTTGTATGTTGATGCTACTCTTT	1475
PT	ribosome entry site (IRES) region, useful in studying HCV replication	QY	784	GC 785	
PT	and expression	DB	1476	GC 1477	
PS	Claim 9; Page -; 69pp; English.				
XX	The invention relates to nucleic acid molecules comprising altered HCV				
XX	NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)				
XX	internal ribosome entry site (IRES) region coding for one or more NS3,				
XX	NS5A, or EMCV IRES mutations, respectively. The location of the				
XX	mutations are detailed in the specification. Also included are				
XX	(1) an expression vector comprising a nucleotide sequence coding for				
XX	the altered nucleic acid, which is transcriptionally coupled to an				
XX	exogenous promoter; (2) a recombinant cell human hepatoma cell comprising				
XX	the altered nucleic acid; (3) a recombinant cell produced by introducing				
XX	into a human hepatoma cell the altered nucleic acid; (4) producing an				
XX	HCV (hepatitis C virus) replicon enhanced cell or which containing a				
XX	functional HCV replicon; (5) an HCV replicon enhanced cells made in the				
XX	method; and (6) measuring the ability of a compound to affect HCV				
XX	activity. The HCV replicons and HCV replicon enhanced cells are useful in				
XX	studying HCV replication and expression, and HCV and host cell				
XX	interactions, producing HCV RNA and proteins, and providing a system				
XX	for measuring the ability of a compound to modulate one or more HCV				
XX	activities e.g. to discover drugs which may treat HCV mediated				
XX	diseases such as liver failure, cirrhosis and hepatocellular carcinoma.				
XX	The present sequence is an HCV replicon Con 1 mutant of the invention.				
XX	Note: The present sequence is not shown in the specification but				
XX	was created by the indexer using the HCV sequence appearing as				
XX	ABK91411 and the information in Claim 9.				
XX	Sequence 9605 BP; 1910 A; 2884 C; 2733 G; 2078 T; 0 other;				
XX	Query Match 88.5%; Score 703.6; DB 24; Length 9605;				
XX	Best Local Similarity 93.7%; Pred. No. 3.4e-179;				
XX	Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;				

Hepatitis C virus Con 1 isolate DNA mutant 9.

HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;

hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;

internal ribosome entry site; IRES; NS5A; HCV replication; mutant.

Hepatitis C virus.

Synthetic.

Mon Dec 22 13:28:52 2003

```
XX PH Key Location/Qualifiers
XX FT CDS 342..9374
XX FT /*tag= a
XX FT /product= "HCV polyprotein"
XX FT /notes= "The polyprotein consists of the Core, E1,
XX FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
XX FT mutation replace (6936,G)
XX FT /*tag= b
XX PN W0200259321-A2.
XX XX 01-AUG-2002.
XX PD 16-JAN-2002; 2002WO-EP00526.
XX PF 23-JAN-2001; 2001US-263479P.
XX PR (RICE-) IST RICE-RIE BIOL MOLECOLARE ANGELETTI.
XX PA De Francesco R, Migliaccio G, Paonessa G;
XX PI WPI; 2002-599793/64.
XX DR New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX PT ribosome entry site (IRES) region, useful in studying HCV replication
XX PT and expression -
XX XX Claim 9; Page -; 69pp; English.
XX XX The invention relates to nucleic acid molecules comprising altered HCV
XX CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX CC internal ribosome entry site (IRES) region coding for one or more NS3,
XX CC NS5A, or EMCV IRES mutations, respectively. The location of the
XX CC mutations are detailed in the specification. Also included are
XX CC (1) an expression vector comprising a nucleotide sequence coupled to an
XX CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
XX CC the altered nucleic acids, which is transcriptionally coupled to an
XX CC the altered nucleic acids; (3) a recombinant cell produced by introducing
XX CC into a human hepatoma cell the altered nucleic acids; (4) producing an
XX CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
XX CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
XX CC method; and (6) measuring the ability of a compound to affect HCV
XX CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
XX CC studying HCV replication and expression, and HCV and host cell
XX CC interactions, producing HCV RNA and proteins, and providing a system
XX CC for measuring the ability of a compound to modulate one or more HCV
XX CC activities e.g. to discover drugs which may treat HCV mediated
XX CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
XX CC The present sequence is an HCV replicon Con 1 mutant of the invention.
XX CC Note: The present sequence is not shown in the specification but
XX CC was created by the indexer using the HCV sequence appearing as
XX CC ABK91411 and the information in Claim 9.
XX SQ Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 other;

Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179; Indels 0; Gaps 0;
Matches 733; Conservative 0; Mismatches 49;

QY 4 TTGGGTAAAGTTCATCGATACCTTACATGCGGGCTTCGCGACCTCGTGGGGTACATCCG 63
DB 696 TTGGGTAAAGTTCATCGATACCTTACATGCGGGCTTCGCGACCTCGTGGGGTACATCCG 755
QY 64 CTCGTGCGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGTCCGGGTCTG 123
DB 756 CTCGTGCGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGTCCGGGTCTG 815
QY 124 GAGGACGGGTGACTATGCAACAGGGAATTCGCGGGTTCCTCTTCTATCTTCCTC 183
DB 816 GAGGACGGGTGACTATGCAACAGGGAATTCGCGGGTTCCTCTTCTATCTTCCTC 875
```

```
QY 184 TTGGCTTTGCTGCTCTGCTGACCGTTCAGCTTCCGCTTATGAAGTGGCGCAACGTGTCC 243
DB |||||
DB 876 TTGGCTTTGCTGCTCTGCTGTTGACCAATCCAGCTTCCGCTTATGAAGTGGCGCAACGTATCC 935
QY 244 GGGATGTACCATGTCAAGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 303
DB |||||
DB 936 GAGGTGTACCATGTCAAGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 995
QY 304 ATGATCATGCACACCCCGGGTGGCTGCGTTCGGGAGAACAACTCTTCCCGCTGC 363
DB |||||
DB 996 ATGATCATGCATACCCCGGGTGGCTGCGTTCGGGAGAACAACTCTTCCCGCTGC 1055
QY 364 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCGAGCGTCCCGCACACGACAATA 423
DB |||||
DB 1056 TGGGTAGCGCTCACTCCGACGCTCGCGGCGGAGAAACGCTAGCGTCCCGCACGATA 1115
QY 424 CGACGCCACGTCGATTTGCTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTGGGG 483
DB |||||
DB 1116 CGACGCCATGTCGATTTGCTGCTGTTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTGGGA 1175
QY 484 GACCTCTGCGGATCTGCTTCTCTGCTCCAGCTGTTCCACATCTCGCTCGCGGCAAT 543
DB |||||
DB 1176 GATCTCTGCGGATCTGCTTCTCTGCTCGTCCGCGGCTGTTCCACCTTCTCGCTCGCGGCAC 1235
QY 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGCGGCGACATACGGGTCAACGTATG 603
DB |||||
DB 1236 GAGACAGTACAGGACTGCAATTTGCTCAATATATCCCGGCGGCGACGTGACGTCAACGTATG 1295
QY 604 CTTTGGGATATGATGAACTGCTGCGCTACAAACGCGCCCTGGTGGTATCGCAGCTGCTC 663
DB |||||
DB 1296 GCTTGGGATATGATGAACTGCTGCGCTACCACTACGACGAGCCCTAGTGGTATCGCAGTTACTC 1355
QY 664 CGGATCCCAAGCTGCTGCGGAGCATGCTGCGGGGGGCCCATTTGGGGAGTCTCGGGGGT 723
DB |||||
DB 1356 CGGATCCCAAGCTGCTGCGGAGCATGCTGCGGGGGGCCCATTTGGGGAGTCTCGGGGGC 1415
QY 724 CTCGCTACTATTCATGCTGGGGAACCTGGGCTAAGGCTTTTGTATTTGTATGCTACTCTTT 783
DB |||||
DB 1416 CTTGCTACTATTCATGCTGGGGAACCTGGGCTAAGGCTTTCTGATTTGTATGCTACTCTTT 1475
QY 784 GC 785
DB 1476 GC 1477
```

Search completed: December 19, 2003, 18:51:03
Job time : 222.785 secs

GenCore version S.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 2045.18 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-5
Perfect score: 795
Sequence: 1 ATGTGGGTAAAGTCATCGA.....TACTCTTTCCTCCTTAATAG 795

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estinu:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hcc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	115.6	14.5	488	9	AV755731
C 2	91.6	11.5	492	9	AV758366
C 3	47.2	5.9	1201	13	BX356664
C 4	43.4	5.5	925	29	CNS0091P

5	42.8	5.4	423	9	AA459034
6	41.6	5.2	1201	9	AL513886
7	41.2	5.2	1195	28	B10902
8	40.6	5.1	402	9	AV392783
9	40.6	5.1	551	9	AV392165
10	40.6	5.1	552	12	B1996341
11	40.6	5.1	584	12	B1727879
12	40.6	5.1	525	10	BE337089
13	40.6	5.0	534	14	CD040840
14	40.6	5.0	671	12	B1723733
15	39.8	5.0	608	14	CB640103
16	39.8	5.0	738	14	CB668031
17	39.8	5.0	969	12	BM017556
18	39.6	5.0	675	12	BQ45314
19	39.2	4.9	571	12	BM692316
20	39.2	4.9	645	29	CNS01213
21	39.2	4.9	1102	13	BQ069967
22	39.2	4.9	359	12	BJ252669
23	39.2	4.9	375	12	BJ246716
24	39.2	4.9	621	14	CAB16001
25	39.2	4.9	624	12	B1723734
26	39.2	4.9	840	29	CC335916
27	39.2	4.9	873	14	CD446071
28	38.8	4.9	590	14	CA659369
29	38.4	4.8	702	14	CD432549
30	38.4	4.8	740	12	BJ536071
31	38.4	4.8	970	29	CNS010C9
32	38.4	4.8	987	29	CNS015VX
33	38.2	4.8	1270	12	BG968359
34	38.2	4.8	354	14	CB966525
35	38.2	4.8	435	14	CB966525
36	37.8	4.8	435	14	C72860
37	37.8	4.8	533	29	CC010084
38	37.8	4.8	659	29	CC405164
39	37.8	4.8	826	29	BZ736582
40	37.8	4.8	895	29	CC359028
41	37.8	4.8	925	29	CC359026
42	37.8	4.8	940	29	CC010085
43	37.8	4.8	951	29	CC405167
44	37.8	4.8	1105	13	BQ064949
45	37.8	4.8	1126	13	BQ063359

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION BM Homo sapiens cDNA clone BMFAK803 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES

primer. Five prime end enriched, double-strand cDNA was

Mon Dec 22 13:28:53 2003

Email: c9apbs-r@mail.nih.gov
This clone is available royalty-free through LBNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 382.

FEATURES

Location/Qualifiers

1. .423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:6032522"
/db_xref="taxon:9606"
/clone="IMAGE:814365"
/tissue_type="germinal center B cell"
/lab_host="DH108"
/clone_lib="NCI_CGAP_GCB1"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTAAGTGGAGCGCGCTCATTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 66 a 162 c 101 g 94 t

Query Match 5.4%; Score 42.8; DB 9; Length 423;
Best Local Similarity 48.7%; Pred. No. 2.3;
Matches 116; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 348 CAACCTCTTCGCGTCTGGTACGCTACCCACAGCTCGACCTAGGACGCCAGCT 407
DB 83 CATGTCGTCTGCTCGCGGAGCCACCGCCCTCCGAGACGAGGAGCTCCCTACGA 142

QY 408 CCCACACGCAATACGACGCGACGTTGCTGCTGGGCGGCTGCTTCTGTC 467
DB 143 CCTAGCGCTCCGCTCCGCGGCGCCATCCCTCTCTCTGCTGCTGCTACTTCTC 202

QY 468 CGCTATGACGTGGGGACCTCTGCGATCTGCTTCTCTGCTCCAGCTGTTACCAT 527
DB 203 CATCAGGAGCAGCTGACTTCAGCGAGTCCCGGAGCACCTGGCTGACAGTTAACAG 262

QY 528 CTGCGCTCCCGCATGACGCGTGCAGCTGCAATGCTCAATCTATCCCGGCCAC 585
DB 263 CAGTCTCTTCCAGCTGAGCCAGCGAGGTTTGGAGGGGCTTCTGCGCCCCCCCAC 320

RESULT 6
AL513886/c 1201 bp mRNA linear EST 08-MAY-2003
LOCUS
DEFINITION
5-PRIME, mRNA sequence.

ACCESSION
AL513886
VERSION
AL513886.2 GI:30463771
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1201)

AUTHORS
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization

JOURNAL
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12777380.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4924.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBA006ZG08RP1&cluster=4924.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA006ZG08RP1.

FEATURES

Location/Qualifiers

1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA006ZG08"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 201 a 311 c 349 g 146 t 194 others

Query Match 5.2%; Score 41.6; DB 9; Length 1201;
Best Local Similarity 26.4%; Pred. No. 6.5;
Matches 116; Conservative 114; Mismatches 206; Indels 3; Gaps 1;

QY 32 GCGGCTTCGCGACCTCGTGGGTACATTCGCTCGCGCCCTAGGCGCGCTG 91
DB 1129 GSGCGCTGCGSBSCTYKGGKSSSSCCSVSSGGGCGSSCCCGGGGGCGGGG 1070

QY 92 CCAGGCGCTGGCGCATGGCGTTCGGGTTCTGGAGACGCGTGAATCATGCAACAGGA 151
DB 1069 CCSCGCGGGGKSSSGSGGCCCGGG---GGSSSSSGSGCGCCMAAASSAYKKGK 1013

QY 152 ATTTCGCGGTTGCTCTTCTCTATCTTCTTCTTCTGCTGCTGCTGCTGACCGTTC 211
DB 1012 GKKTCTTTTMAAATWTHTTTTWTCTTAAGGGGTAKCAKCCCMCCCGGCTGS 953

QY 212 CAGCTTCGCTTATGAAGTGGCGCAACGTTCCGGATGTACCATGTCAAGACGACTGCT 271
DB 952 GACSCSCCAADAACVCGAGMDSGAMKSGTGVSCCTTSRCKWGGTTSGMMVGCATTYM 893

QY 272 CCAACTCAAGCATTTGTATGAGGACGCGGACATGATCATGCACACCCCGGTCGTCG 331
DB 892 AYBSYTGTRWTTTGTGCTYASGSGSYSSRKBGKCCMAYACSCGAGASCST 833

QY 332 CTGCGGTTGGGAGAACAACTCTTCCGCTGCTGGTAGCGCTCACCCCGGCTGCGCAG 391
DB 832 SGCSKGTNTTNTTGTCTGTTGAAGASMBGRTWAGGGGGGGGGCCGCCMCCCCCYB 773

QY 392 CTAGGAACGCGACGCTCCCCACCCACGACCAATACGACCGCATGCTGCTGCTGGGG 451
DB 772 BBCCMCHCTKCSKWCRCGACTYCCCASSCTSYGGTCCCTTCTTGTGTTBSH 713

QY 452 CGGCTGCTTCTGTTCCGC 470
DB 712 TKTTTTTTTCTTCCCVSC 694

RESULT 7

BI0902 1195 bp DNA linear GSS 14-MAY-1997
LOCUS
DEFINITION
F13A13-Sp6 10F Arabidopsis thaliana genomic clone F13A13, genomic survey sequence.

ACCESSION
BI0902
VERSION
BI0902.1 GI:2092024

KEYWORDS
GSS.
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

Internet: <http://www.kazusa.or.jp/en/plant/>.

FEATURES

```
1 551
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM083e05_r"
/dev_stage="photoautotrophic growth"
/clone_lib="Chlamydomonas reinhardtii C9"
/note="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
```

BASE COUNT	94 a	182 c	189 g	85 t	1 others
ORIGIN					

Query Match	5.1%;	Score 40.6;	DB 9;	Length 551;
Best Local Similarity	45.3%;	Pred. No. 9.1;	Indels	Gaps
Matches	148;	Conservative	0;	Mismatches 179;
QY	42	CGACCTCGTGGGTACATTCGCGTTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGAGCCCT	101	
DB	108	CGAGCTATCTCGTTCATTTGTCGGCGCACTGCCAAATGAAGAGCGTGTGACGAGCT	167	
QY	102	GGGCGCATGGCGTCCGGTTCCTCGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGG	161	
DB	168	GGCGCGCGCGCGAGTGGAGGGCGGCTACGCGCAGATTCGTCGTCGTGAGCTTTGGGCG	227	
QY	162	TTGTCTCTTTCTCTATCTTCTCTTGGCTTTGCTGTCTGTCTGACCGTTTCCAGCTTTCGCG	221	
DB	228	CCGCAAGGTGTTTGAACAGATCAAGAGTACGTGTGAACTCAAGSCCCAGAACCCAG	287	
QY	222	TTATGAAGTGGCAACGTGTTCGGGATGTACCATGTACGAACGATGTCCTCAACTCAAG	281	
DB	288	CTTCGCGCTCGCTGCTGGGCCACTCGCTGGCGCGCGCACCGCGCGCTGCCTGTCGAT	347	
QY	282	CATTGTGTATGAGCAGCGAGCATGATCATGACACACCCCGGCGCTGTCGCTGCTCG	341	
DB	348	CCTGATGCACCAACAGCAGGAGTTTCGGCGCGCATCTACGGCGCGGTGCCCATGCCGG	407	
QY	342	GGAGAACACTCTTCCTCCGCTGCTGGGT	368	
DB	408	CAAGAAGAGCAAGGGCAGCTACATGAT	434	

RESULT 10
B1996341
LOCUS
DEFINITION
Cc.reinhardtii CG-1690, Stress II (normalized), C. reinhardtii cDNA, mRNA sequence.
linear EST 25-OCT-2001 mRNA 552 bp
B1996341

ACCESSION	BI9996341	GI:16431115
REFSEQ	BT9996341.1	

EST.
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 552)
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
, P., McDermott, J. P., Strager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031

JOURNAL
Unpublished

COMMENT

DCMB Box 91000

Duke University
Durham NC 27708-1000

Durham, NC 27708-1000
Tel: 919 613 8159

Tel: 919 613 8177
Fax: 919 613 8177

Email: chauser@duke.edu

FEATURES	Location/Quali
----------	----------------

```
source
1. .552
/*****cm="ch]
```

считается, что

```

/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clo_nlb="C. reinhardtii CC-1690, Stress II (normalized)
, lambda Zap II"
/notes="vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoRI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."

```

BASE COUNT	93 a	184 c	189 g	86 t	
ORIGIN					
Query Match	5.1%	Score 40.6;	DB 12;	Length 552;	
Best Local Similarity	45.3%;	Pred. NO. 9.1;			
Matches 148;	Conservative 0;	Mismatches 179;	Indels 0;	Gaps 0;	
QY	42	CGACCTCGTGGGGTACATTCGCTGTGGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCT	101		
Db	110	CGAGTTCATCTGTGTCATTTGTGGCGGCACATGCCAAATGAAGAGCTGTCTACGGACCT	169		
QY	102	GGCGATGGCGTCCGGGTTCTGGAGGAGCGCGTGAACATATGCAACAGGGAATTTGCCCGG	161		
Db	170	GGCGCGCGCGGCGCGAGTGGGAGGCGGCTACGGCGCAGACTCGTGAGCTTTGGGCGC	229		
QY	162	TTGCTCTTTCTATCTTCTCTTGGCTTTTGCTGTCTGTGTGACCGTTCCAGCTTCCGC	221		
Db	230	CGCAAGGTGTTTGACGAGATCAAGAGTAGCTGTGAACCTCAAGGCCCAGAACCCGAG	289		
QY	222	TTATGAAGTGCACACGTGTCCGGGATGTACCATGTACGAACGACTGCTTCAACTCAAG	281		
Db	290	CTTGGCCGTCGCTGGTGGGCGCACTCGCTGGGCGGCGCACCGCGGCTGCCCTGTGAT	349		
QY	282	CATTGTGTATGAGGACGGGACATGATCATGCACACCCGGGTGGGTGCCCTTCGCTTGC	341		
Db	350	CCTGATGCACACGACGAGGAGTTTGGCGGCGCGCATCTACGGGCGGTGCCCATGCCGG	409		
QY	342	GGAGAACAACTCTTCCGCTGCTGGGT	368		
Db	410	CAGAAGAGCAAGGCGAGCTACATGAT	436		

DESMITT 11

REF ID: A61727879

LOCUS

DEFINITION

1

ACCESSION
NUMBER

VERSION
KEYWORDS

REFERENCES

ORGANIS

REFERENCE

AUTHORS

TITLE

JOURNAL
COMPTON

COMMENT

DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

Location/Qualifiers
1. .584
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 106 a 188 c 197 g 93 t

ORIGIN

Query Match 5.1%; Score 40.6; DB 12; Length 584;
Best Local Similarity 45.3%; Pred. No. 9.3;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGTACATTCGCTGCTGGCGCCCTAGGGGGCGCTCCAGGGCCCT 101
|||
Db 47 CGAGCTCATCTGCTGCTGGCGCACTGCCAATCAAGGAGCTGCTGACGACCT 106
|||
QY 102 GCGCATGGCTCGGGTCTGAGAGCGCGTGAATATGCAACAGGAGATTTGCCGG 161
|||
Db 107 GCGCGCGCGCGCGAGTGGAGGGCGGCTACGCGACGAGTCCGTGAGCTTGGCGC 166
|||
QY 162 TTCTCTTCTATCTTCTCTGCTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 221
|||
Db 167 CCGCAAGTGTGTCAGAGATCAAGAGTACGTGCTGAACTCAAGCCAGAACCCAG 226
|||
QY 222 TTATGAAGTGCACAGCTGTCGGGATGTACCATGTACGAACGAGTGTCCAACTCAAG 281
|||
Db 227 CTTGCGCGTCCGTGCTGGGCACTCGCTGGCGCGGCGCACCGCGCTGCTGCTGAT 286
|||
QY 282 CATTGTGTAGCGAGCGACATGATCATGACACCCCGGTGCGTGCCTGCGTTGCG 341
|||
Db 287 CTTGATGACACGACGAGAGTTTGGCGCGCATCTACGGGCGGCTGCTCCATGCGGG 346
|||
QY 342 GGAGAACACTCTTCCCGCTGCTGGGT 368
|||
Db 347 CAAGAGAGCAAGGGAGCTACATGAT 373
|||

RESULT 12
LOCUS BE337089 525 bp mRNA linear EST 14-JUL-2000
DEFINITION 894043G08.y1 C. reinhardtii CC-1690, normalized, Lambda Zap II
Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BE337089
VERSION BE337089.1 GI:9210174
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 525)
Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
Unpublished
Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu.

Location/Qualifiers
1. .525
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light. TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 99 a 181 c 160 g 85 t

ORIGIN

Query Match 5.0%; Score 40; DB 10; Length 525;
Best Local Similarity 47.8%; Pred. No. 13;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 261 GAACGACTGTCTCAACTCAAGCATTTGTATAGAGGAGGAGCATGATCATGACACCCC 320
|||
Db 43 GCACCGCTTCAACAACCCACCGTGTGAATTCGCCGCTACAAGTACCTGTGTCAT 102
|||
QY 321 CGGGTGGTGGCTCGTTCGGGAGAACAACTTTCCCGCTGTGGGTAGCGGTACACCC 380
|||
Db 103 GGAGGACAAGACCTGCCACCTGGAGGTCATCGAGTACTGCGACCTGGGCAACCTGTCAA 162
|||
QY 381 CACGCTCGGAGCTAGGACGCCAGCGTCCCAACACACACACACACACACACACGTCGATTT 440
|||
Db 163 CGCGCTCAAAAACAAACATCTTCATGATCCCCCAACCCCGTCATCGCGCGCGCGGCGC 222
|||
QY 441 GCTCGTTGGCGCGGCTGCTTTCTGTTCCGCTATGTAGTGGGGAGACCTCTCGCGATCTGT 500
|||
Db 223 GGGCGAGCGCGCGCGCGGCGGAGCTAGGGAGCGCGCGGCGGAGCCATGAGGT 282
|||
QY 501 CTTCTCGTCTCCAGCTGTTTAC 524
|||
Db 283 CAACATGCGCACCTGCTGCTCAC 306
|||

RESULT 13
LOCUS CD040840 534 bp mRNA linear EST 09-MAY-2003
DEFINITION CD040840
peHB036xB09f 300663 peHB: Infected hypocotyl soybean host. 48 hrs
post infection Phytophthora sojae cDNA clone SHB036B09 5, mRNA
sequence.
ACCESSION CD040840
VERSION CD040840.1 GI:30502701
KEYWORDS EST.

[illegible]

Search completed: December 20, 2003, 06:54:39
Job time : 2050.18 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 1828.77 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-9

Perfect score: 483

Sequence: 1 ATGCCCGGTGCTCTTCTC.....TGATGAAGTGGTCTTAATAG 483

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 288711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_cm.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pi.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rnd.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	483	100.0	483	6	A48671	A48671 Sequence 9
2	483	100.0	483	6	AR157327	AR157327 Sequence
3	483	100.0	483	6	AX452758	AX452758 Sequence
4	483	100.0	483	6	AX685010	AX685010 Sequence
5	474.2	98.2	480	6	A48673	A48673 Sequence 11
6	474.2	98.2	480	6	AR157328	AR157328 Sequence
7	474.2	98.2	480	6	AX452760	AX452760 Sequence
8	474.2	98.2	480	6	AX685012	AX685012 Sequence
9	438.2	90.7	9379	14	AF207766	AF207766 Hepatitis C
10	435	90.1	1880	14	HPC5TRJ4	D00832 Hepatitis C
11	435	90.1	2540	6	E04260	E04260 cDNA encodi
12	435	90.1	2540	6	E04805	E04805 cDNA to 5'-
13	435	90.1	2540	6	E07391	E07391 cDNA encodi
14	435	90.1	9448	14	HPCJ483	D13558 Hepatitis C
15	433.4	89.7	1539	6	AR027786	AR027786 Sequence
16	433.4	89.7	1863	6	AR027783	AR027783 Sequence
17	433	89.6	9580	14	AF054250	AF054250 Hepatitis
18	431.8	89.4	8780	14	AF054257	AF054257 Hepatitis
19	431.8	89.4	8780	14	AF054258	AF054258 Hepatitis
20	430.2	89.1	8780	14	AF054255	AF054255 Hepatitis
21	430.2	89.1	8781	14	AF054254	AF054254 Hepatitis
22	430.2	89.1	9379	14	AF165052	AF165052 Hepatitis
23	430.2	89.1	9460	14	HPCJ491	D10750 Hepatitis C
24	430.2	89.1	9595	6	AR119832	AR119832 Sequence
25	430.2	89.1	9595	14	AF054247	AF054247 Hepatitis
26	430.2	89.1	9596	14	AF054249	AF054249 Hepatitis
27	430.2	89.1	9599	6	AR119833	AR119833 Sequence
28	428.6	88.7	577	6	E04085	E04085 gDNA encodi
29	428.6	88.7	618	14	HPCJ1F12	D28929 Hepatitis C
30	428.6	88.7	8779	14	AF054251	AF054251 Hepatitis
31	428.6	88.7	8780	14	AF054252	AF054252 Hepatitis
32	428.6	88.7	9377	14	AF207756	AF207756 Hepatitis
33	428.6	88.7	9410	14	HPCJ1R2	D50481 Hepatitis C
34	428.6	88.7	9435	14	AB049093	AB049093 Hepatitis
35	428.6	88.7	9595	14	AF054248	AF054248 Hepatitis
36	427	88.4	1595	14	HPCNSISPF	M74809 Hepatitis C
37	427	88.4	1635	14	HPCNSISPD	M74807 Hepatitis C
38	427	88.4	3296	14	AB008446	AB008446 Hepatitis
39	427	88.4	8780	14	AF054259	AF054259 Hepatitis
40	427	88.4	8781	14	AF054256	AF054256 Hepatitis
41	427	88.4	9361	14	AF483269	AF483269 Hepatitis
42	427	88.4	9369	14	AF165054	AF165054 Hepatitis
43	427	88.4	9379	14	AF165051	AF165051 Hepatitis
44	425.4	88.1	577	6	E04089	E04089 gDNA encodi
45	425.4	88.1	1413	14	HPCRNAP	D00574 Hepatitis C

ALIGNMENTS

RESULT 1
A48671
LOCUS A48671 Sequence 9 from Patent WO9604385.
DEFINITION A48671
ACCESSION A48671
VERSION A48671.1 GI:2302384
KEYWORDS .
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 483)
Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 9 15-FEB-1996;

[illegible]

[illegible]

REFERENCE	1 (bases 1 to 2540)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
-----------	---------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

[illegible]

Db	806	CTGCTTTCTGCTCCGCTATGTACGTGGGGGATCTCGGGATCTGTTTTCCTCGTCC	865
Qy	362	AGTGTTCACTTTCACCTCCGGGATCAACAGTACAGGACTCAACTGCTCAATCT	421
Db	866	AGCTGTTCACTTCTCGGCTCGCGGATGAGACAGTGCAGGACTGCAACTGCTCAATCT	925
Qy	422	ATCCGGCCATGATATCAGGTCAACCGCATGGCTGGGATATGATGATGAATGTC	476
Db	926	ATCCGGCCATTTATCAGGTACCCGATGGCTTGGGATATGATGATGAATGGTC	980

Search completed: December 20, 2003, 02:01:52
Job time : 1830.77 secs

QY	122	ACTCAAGCATAGTGTATGAGGACGGGACATGATCATGACACACCCCGGGTGGCGCCT	181
DB	967	ACTCAAGCATTTGTATGAGGACGGGACATGATCATGACATCTCCGGGTGGCGCCT	1026
QY	182	GGCTTGGGAGGCAACTCCTCCCGTTGCTGGGTGGCGGTCACTCCACGCTCGGGCCA	241
DB	1027	GGTTTGGGAGGACACAGCTCCCGTTGCTGGGTAGCGTCACTCCACGCTCGGGCCA	1086
QY	242	GGAAAGCCAGCGTCCCAACAGCAATACGACGCGCATCAAGTTCGTTGGGGGTG	301
DB	1087	GGAAAGCCAGCGTCCCAACAGCAATACGACGCGCATCAAGTTCGTTGGGGGTG	1146
QY	302	CTGCTTCTTGTTCGGCTATGTAGTGCGGGGATCTTCGGGATCTGTTTCTGTTTCCC	361
DB	1147	CTGCTTCTTGTTCGGCTATGTAGTGCGGGGATCTTCGGGATCTGTTTCTGTTTCCC	1206
QY	362	AGCTGTTCACTTCTCACTTCGCGGATCAACAGTACAGGACTGCAACTGCTCAATCT	421
DB	1207	AGCTGTTCACTTCTCACTTCGCGGATCAACAGTACAGGACTGCAACTGCTCAATCT	1266
QY	422	ATCCGGCCATGTATCAGGTCAACGCATAGCTTGGGATATGATGAATGTC	476
DB	1267	ATCCGGCCATTTATCAGGTCAACGCATAGCTTGGGATATGATGAATGTC	1321
RESULT 15			
AR027786			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 133.529 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303a-9
Perfect score: 483
Sequence: 1 ATGCCCGTGTCTTCTC.....TGATGAACGTGCTCTAATAG 483

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	483	100.0	483	17	AAT12707
2	483	100.0	483	24	AA148916
3	474.2	98.2	480	17	AAT12708
4	474.2	98.2	480	24	AA148917
5	435	90.1	1880	13	AAQ24467
6	435	90.1	2187	19	ABAQ3491
7	435	90.1	2540	14	AAQ43889
8	435	90.1	2540	15	AAQ63753

9	433.4	89.7	1863	12	AAQ15363	Fragment of NANB h
10	431.8	89.4	2540	13	AAQ29628	Hepatitis C virus
11	430.2	89.1	9595	20	AAQ24843	Infectious hepatitis
12	430.2	89.1	9595	22	AAC86939	Nucleotide sequenc
13	430.2	89.1	9595	22	AAC23492	Infectious hepatitis
14	430.2	89.1	9599	20	AAQ24833	Infectious hepatitis
15	428.6	88.7	577	14	AAQ35081	HCV envelope regio
16	428.6	88.7	2187	19	ABAQ3492	Cuticle protein 1
17	425.8	88.2	580	12	AAQ11076	Fragment of hepati
18	425.8	88.2	580	20	AAQ207647	HCV J1 E domain co
19	425.4	88.1	577	14	AAQ35085	HCV envelope regio
20	425.4	88.1	577	16	AAQ79750	Hepatitis C virus
21	425.4	88.1	580	20	AAQ26733	Consensus sequence
22	425.4	88.1	580	20	AAQ26728	Consensus sequence
23	425.4	88.1	580	20	AAQ00401	Hepatitis C virus
24	425.4	88.1	1249	16	AAQ79772	Hepatitis C virus
25	425.4	88.1	1249	20	AAQ26739	Consensus sequence
26	425.4	88.1	1562	19	AAV60672	Fragment #5 isolat
27	425.4	88.1	1953	25	AAU55222	Plasmid pDK2 DNA
28	425.4	88.1	2829	19	AAV60673	Fragment #6 isolat
29	425.4	88.1	3401	15	AAQ64069	Non-A, non-B hepat
30	425.4	88.1	3401	16	AAQ30387	5'UTR/CORE/ENV/NS1
31	425.4	88.1	3461	15	AAQ64068	Non-A, non-B hepat
32	425.4	88.1	3461	16	AAQ30386	5'UTR/CORE/ENV/NS1
33	425.2	88.0	567	13	AAQ27160	NANB hepatitis vir
34	424.8	88.0	642	17	AAT12704	HCV E1 construct H
35	424.8	88.0	642	24	AA148913	Hepatitis C virus
36	424.4	87.9	1270	19	AAV60668	Fragment #1 isolat
37	423.8	87.7	577	14	AAQ35086	HCV envelope regio
38	423.8	87.7	580	16	AAQ79758	Hepatitis C virus
39	423.8	87.7	580	16	AAQ79759	Hepatitis C virus
40	423.8	87.7	580	16	AAQ79760	Hepatitis C virus
41	423.8	87.7	795	17	AAT12705	HCV E1 construct H
42	423.8	87.7	795	24	AA148914	Hepatitis C virus
43	423.8	87.7	1249	20	AAQ00460	Hepatitis C virus
44	423.8	87.7	1682	13	AAQ27159	Hepatitis C virus
45	423.8	87.7	2082	24	AA148939	NANB hepatitis C virus

ALIGNMENTS

RESULT 1
AAT12707
ID AAT12707 standard; DNA; 483 BP.

AC AAT12707;
XX
XX
DT 23-SEP-1996 (first entry)
XX
XX
DE HCV E1 construct HCC112A.

KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.

OS Hepatitis C virus.

PN WQ9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP03031.

XX 29-JUL-1994; 94EP-0870132.

XX (INNO-) INNOGENETICS NV.

PI Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

proteins - in presence of di:sulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX
 XX Claim 23; Fig 21; 146pp; English.
 PS
 XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a
 CC disulphide bond cleavage agent, after lysis of recombinant host cells.
 CC The constructs containing the purified HCV envelope proteins can be used
 CC for vaccinating humans against HCV, for in vitro detection of HCV
 CC antibodies in a sample, and in a serotyping assay for detecting one or
 CC more serological types of HCV present in a biological sample. The
 CC constructs can also be immobilised on a solid substrate and incorporated
 CC into a reversed phase hybridisation assay for determining the presence or
 CC the genotype of HCV. The new purification method preserves the
 CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
 CC eliminates contaminating proteins. Antigens isolated using this method
 CC are more reactive with human sera than those isolated by known
 CC techniques.
 XX
 XX Sequence 483 BP; 85 A; 152 C; 123 G; 123 T; 0 other;
 SQ
 Query Match 100.0%; Score 483; DB 17; Length 483;
 PT Best Local Similarity 100.0%; Pred. No. 2.6e-131; Indels 0; Gaps 0;
 PT Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 ATGCCCGGTGCTCTTTCTATCTCTCTTGGCCCTGCTGCTCTGTCGACCATACCA 60
 Db 1 ATGCCCGGTGCTCTTTCTATCTCTCTTGGCCCTGCTGCTCTGTCGACCATACCA 60
 QY 61 GCTTCCGCTTATGAAGTGGCAACGTGTCGGGGTGACCATGTGTCGACCATGCTGCC 120
 Db 61 GCTTCCGCTTATGAAGTGGCAACGTGTCGGGGTGACCATGTGTCGACCATGCTGCC 120
 QY 121 AACTCAAGCATATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 Db 121 AACTCAAGCATATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 QY 181 TGCGTTCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 Db 181 TGCGTTCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 QY 241 AGGAACGCCAGGCTCCCAACAGCAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 Db 241 AGGAACGCCAGGCTCCCAACAGCAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 QY 301 GCTGCTTTCTGTTCCGCTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 Db 301 GCTGCTTTCTGTTCCGCTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 QY 361 CAGCTGTTCACCTTCTCACCCTCGCGGATCTCTGCGGATCTCTGCGGATCTCTGCGG 420
 Db 361 CAGCTGTTCACCTTCTCACCCTCGCGGATCTCTGCGGATCTCTGCGGATCTCTGCGG 420
 QY 421 TATCCCGGCGCATGATCAGGTTCACCGCATGCGGATGATGATGATGATGATGATGATGAT 480
 Db 421 TATCCCGGCGCATGATCAGGTTCACCGCATGCGGATGATGATGATGATGATGATGATGAT 480
 QY 481 TAG 483
 Db 481 TAG 483

RESULT 2
 ID AAL48916
 XX AAL48916 standard; DNA; 483 BP.
 AC AAL48916;

XX 24-OCT-2002 (first entry)
 XX Hepatitis C virus clone HCC112A E1 protein coding sequence.
 XX
 XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 XX virucide; immunostimulant; vaccine; ds.
 XX
 XX Hepatitis C virus.
 XX WO200255548-A2.
 XX 18-JUL-2002.
 XX 11-JAN-2002; 2002WO-EP00219.
 XX 11-JAN-2001; 2001US-260699P.
 XX 30-AUG-2001; 2001US-315768P.
 XX (INNO-) INNOGENETICS NV.
 XX Maertens G, Bosman F, Buyse M;
 XX WPI; 2002-599657/64.
 XX P-PSDB; AAO18663.
 XX
 XX New therapeutic vaccine compositions comprising at least one purified
 XX recombinant hepatitis C virus (HCV) single or specific oligomeric
 XX recombinant envelope protein E1 or E2, useful for immunizing humans
 XX from HCV infection -
 XX Example 2; Page 165-166; 243pp; English.
 XX
 XX The present invention relates to new therapeutic vaccine compositions for
 XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 XX composition containing at least one purified recombinant HCV single or
 XX specific oligomeric recombinant envelope proteins selected from an E1 and
 XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 XX useful for inducing HCV-specific antibodies or for immunising humans
 XX against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 XX vaccines or therapeutics, in HCV screening and confirmatory antibody
 XX tests, for raising antibodies, in the preparation of medicament, and for
 XX in vitro monitoring of HCV disease or prognosis of the response to
 XX treatment of patients suffering from HCV infection. The present sequence
 XX is a coding sequence described in the exemplification of the invention.
 XX
 XX Sequence 483 BP; 85 A; 152 C; 123 G; 123 T; 0 other;
 SQ
 Query Match 100.0%; Score 483; DB 24; Length 483;
 PT Best Local Similarity 100.0%; Pred. No. 2.6e-131;
 PT Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 ATGCCCGGTGCTCTTTCTATCTCTTGGCCCTGCTGCTCTGTCGACCATACCA 60
 Db 1 ATGCCCGGTGCTCTTTCTATCTCTTGGCCCTGCTGCTCTGTCGACCATACCA 60
 QY 61 GCTTCCGCTTATGAAGTGGCAACGTGTCGGGGTGACCATGTGTCGACCATGCTGCC 120
 Db 61 GCTTCCGCTTATGAAGTGGCAACGTGTCGGGGTGACCATGTGTCGACCATGCTGCC 120
 QY 121 AACTCAAGCATATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 Db 121 AACTCAAGCATATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 QY 181 TGCGTTCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 Db 181 TGCGTTCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 QY 241 AGGAACGCCAGGCTCCCAACAGCAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 Db 241 AGGAACGCCAGGCTCCCAACAGCAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 QY 301 GCTGCTTTCTGTTCCGCTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 Db 301 GCTGCTTTCTGTTCCGCTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360

[illegible]

CC NANB hepatitis virus strain HC-J4 was isolated from a plasma sample
CC of a chimpanzee challenged with NANB hepatitis for infectivity but
CC which tested negative for HCV antibody by Ortho HCV Ab ELISA test.
CC RNA was isolated from the sample and reverse transcribed into cDNA.
CC The 513 amino acids encoded by the CDS were determined but are not
CC given in the specification (and hence are not included in A-Geneseq).
CC A study of the deduced sequence suggested that the CDS encodes NANBH
CC virus core proteins. Primers for detecting NANB hepatitis virus were
CC designed based on the HC-J4 sequence.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 1863 BP; 333 A; 586 C; 547 G; 397 T; 0 other;

Query Match 89.7%; Score 433.4; DB 12; Length 1863;
Best Local Similarity 94.5%; Pred. No. 1.4e-116;
Matches 449; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCCCTGCTGCTGCTGACCATACCAG 61
Db |||||
QY 830 TGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTTGTGCTGTTGACCATCCAG 889
Db |||||
QY 62 CTTCCGCTTATGAAGTGCACAGCTGTCCGGGTGTACCATGTACAGAACGACTGTCTCCA 121
Db |||||
QY 890 CTTCCGCTTATGAAGTGCACAGCTGTCCGGGTGTACCATGTACAGAACGACTGTCTCCA 949
Db |||||
QY 122 ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGCACACCCCGGTGTGCGCCT 181
Db |||||
QY 950 ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGCATCTCCCGGTGTGCGCCT 1009
Db |||||
QY 182 GCGTTCCGGAGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGCCA 241
Db |||||
QY 1010 GCGTTCCGGAGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGCCA 1069
Db |||||
QY 242 GGAACGCGACGTCCTCCACAAACGACATACAGCCACGTCGATTTGCTCGTTGGGCTG 301
Db |||||
QY 1070 GGAATGCCAGGTCCTCCACAAACGACATACAGCCACGTCGATTTGCTCGTTGGGCGG 1129
Db |||||
QY 302 CTGCTTTCTGTCGCTATGATGATGAGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
Db |||||
QY 1130 CTGCTTTCTGTCGCTATGATGATGAGGATCTCTGCGGATCTGTTTCTGTTTCCC 1189
Db |||||
QY 362 AGCTGTTCACTTCTCACTCCGCGCATCAACAGTACAGGACTGCACTGCTCAATCT 421
Db |||||
QY 1190 AGCTGTTCACTTCTCGCTCGCGCATGAGACAGTGCAGGACTGCACTGCTCAATCT 1249
Db |||||
QY 422 ATCCCGGCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGAATGGTTC 476
Db |||||
QY 1250 ATCCCGGCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGAATGGTTC 1304
Db |||||

RESULT 10
AAQ29628
ID AAQ29628 standard; DNA; 2540 BP.
XX
AC AAQ29628;
XX
DT 25-MAR-2003 (updated)
DT 16-MAR-1993 (first entry)
DE Hepatitis C virus HC-J4 5' region.
XX
KW Non-A non-B hepatitis; NANBH; HCV; detection; diagnosis; screening;
KW PCR; primer; polymerase chain reaction; ss.
XX
OS Hepatitis C virus.
XX
PN EP510952-A1.
XX
PD 28-OCT-1992.
XX
PF 23-APR-1992; 92EP-0303625.
XX

PR 26-APR-1991; 91JP-0191376.
XX
PA (IMMO) IMMUNO JAPAN INC.
XX
PI Nakamura T, Okamoto H;
XX
XX WPI; 1992-359137/44.
XX
PT Detection of non-A, non-B hepatitis virus - using new
PT oligo-nucleotide primers with nucleotide sequences corresp. to
PT part. of the viral RNA
XX
XX
PS Disclosure; Page 18; 54pp; English.
XX
CC This sequence represents the 5' region of hepatitis C virus RNA. The
CC original sample was obtained from human and chimpanzee plasma. RNA
CC was isolated from several samples and homology compared, and the
CC respective sequence of about 1900 - 2500 nucleotides of the 5'
CC terminus and 1100 nucleotides of the 3' terminus determined. The 5'
CC region (given) contains a non-coding region of at least 340
CC nucleotides and a region coding for the structural protein followed
CC by a region coding for the non-structural protein (none actually
CC detailed on the sequence given in the specification). When compared
CC with the sequence of HCV disclosed in EP-388232 this sequence showed
CC homology of 80.5%.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2540 BP; 472 A; 775 C; 741 G; 552 T; 0 other;

Query Match 89.4%; Score 431.8; DB 13; Length 2540;
Best Local Similarity 94.3%; Pred. No. 4.5e-116;
Matches 448; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTTCTCTATCTTCTTGGCCCTGCTGCTGCTGACCATACCAG 61
Db |||||
QY 847 TGCCCGGTTGCTCTTTCTCTATCTTCTCATGCTTGTGCTGTTGACCATCCAG 906
Db |||||
QY 62 CTTCCGCTTATGAAGTGCACAGCTGTCCGGGTGTACCATGTACAGAACGACTGTCTCCA 121
Db |||||
QY 907 CTTCCGCTTATGAAGTGCACAGCTGTCCGGGTGTACCATGTACAGAACGACTGTCTCCA 966
Db |||||
QY 122 ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGCACACCCCGGTGTGCGCCT 181
Db |||||
QY 967 ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGCATCTCCCGGTGTGCGCCT 1026
Db |||||
QY 182 GCGTTCCGGAGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGCCA 241
Db |||||
QY 1027 GCGTTCCGGAGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGCCA 1086
Db |||||
QY 242 GGAACGCGACGTCCTCCACAAACGACATACAGCCACGTCGATTTGCTCGTTGGGCTG 301
Db |||||
QY 1087 GGAATGCCAGGTCCTCCACAAACGACATACAGCCACGTCGATTTGCTCGTTGGGCGG 1146
Db |||||
QY 302 CTGCTTTCTGTCGCTATGATGATGAGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
Db |||||
QY 1147 CTGCTTTCTGTCGCTATGATGATGAGGATCTCTGCGGATCTGTTTCTGTTTCCC 1206
Db |||||
QY 362 AGCTGTTCACTTCTCACTCCGCGCATCAACAGTACAGGACTGCACTGCTCAATCT 421
Db |||||
QY 1207 AGCTGTTCACTTCTCGCTCGCGCATGAGACAGTGCAGGACTGCACTGCTCAATCT 1266
Db |||||
QY 422 ATCCCGGCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGAATGGTTC 476
Db |||||
QY 1267 ATCCCGGCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGAATGGTTC 1321
Db |||||

RESULT 11
AAQ24843
ID AAQ24843 standard; DNA; 9595 BP.
XX
XX
AC AAQ24843;
XX
DT 21-JUN-1999 (first entry)

XX	Infectious hepatitis C virus genotype 1b strain HC-J4 genome.
DE	HCV; infectious clone; infection; diagnosis; therapy; vaccine;
KW	screening; assay; antiviral; virucide; ss.
KM	
XX	Hepatitis C virus.
OS	
XX	
FH	Key Location/Qualifiers
FT	CDS 342..9374
FT	/*tag= a
XX	
PN	WO9904008-A2.
XX	
XX	28-JAN-1999.
XX	16-JUL-1998; 98WO-US14688.
PF	
XX	27-JAN-1998; 98US-0014416.
PR	18-JUL-1997; 97US-0053062.
XX	
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PA	
XX	Bukh J, Emerson SU, Purcell RH, Yanagi M;
PI	
XX	WPI; 1999-132252/11.
DR	P-PSDB; AAW98022.
XX	
XX	New isolated hepatitis C virus nucleic acids - used to develop
PT	products for the diagnosis, prevention and treatment of HCV
PT	infections and for developing screening assays
PT	
XX	Claim 3; Fig 14A-F; 126pp; English.
PS	
XX	The present sequence comprises the nucleic acid sequence of the
CC	genome of infectious hepatitis C virus (HCV) genotype 1b strain
CC	HC-J4 (ATCC 209596) that is capable of expressing this virus when
CC	transfected into cells. HC-J4 was obtained from acute phase plasma
CC	of a chimpanzee experimentally infected with serum containing
CC	HC-J4/91. The claimed infectious nucleic acid sequence can be used
CC	to produce chimeric genomes (see AMX24833) consisting of the open
CC	reading frames of infectious nucleic acid sequences of other
CC	genotypes (including genotypes 1-6) and subtypes (such as 1b, 2a,
CC	2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also relates to
CC	the introduction of mutations or deletions into infectious nucleic
CC	acid sequences in order to produce an attenuated HCV virus suitable
CC	for vaccine development. Infectious nucleic acid sequences can
CC	also be used to produce attenuated virus via passage in vitro or in
CC	vivo of the viruses produced by transfection of a host cell with
CC	the infectious nucleic acid sequence. Vaccines comprising one or
CC	more polypeptides made from the infectious nucleic acid sequence are
CC	used to immunise mammals, especially humans, against hepatitis C.
CC	The nucleic acid sequences can also be used to induce protective
CC	immunity against the virus. The nucleic acid sequences or their
CC	encoded proteases (e.g. NS3 protease) can additionally be used to
CC	develop screening assays to identify antiviral agents for HCV.
XX	
SQ	Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;
Query Match 89.1%; Score 430.2; DB 20; Length 9595;	
Best Local Similarity 94.1%; Pred. No. 2.1e-115;	
Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;	
QY	2 TGCCCGGTGCTCTTTCTCTATCTTCTTGCCCTGTCCTGCTGCACCATCCAG 61
Db	847 TGCCCGGTGCTCTTTCTCTATCTTCTTGCCCTGTCCTGCTGCACCATCCAG 906
QY	62 CTTCGGCTTATGAAGTCGCAACGCTGTCGGGGTGTPACCATGTCACGACGACTCTCCA 121
Db	907 CTTCGGCTTATGAAGTCGCAACGCTGTCGGGGTATACCATGTCACGACGACTCTCCA 966
QY	122 ACTCAACGATAGTGTATGAGCGCGGACATGATGATGACACCCCCGGTGCCTCCT 181

QY	182	GCGTTGGGAGGCAACTCTCCGTTTGGTGGGTCGCTCACTCCACGCTCGCGGCCA	241
Db	1027	GTGTTTCAGGAGGGTAACAGCTCCCGTTGCTGGGTAGGCTCACTCCACGCTCGCGGCCA	1086
QY	242	GGAACGCCAGGCTCCCACAACGACAATACGACGCCACGTCGATTGCTGTTGGGCTG	301
Db	1087	GGAATGCCAGCGTCCCCTAGCACAAATACGACGCCACGTCGATTGTTGGGACGG	1146
QY	302	CTGCTTTCTGTTCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	361
Db	1147	CTGCTTTCTGTTCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1206
QY	362	AGCTGTTTCACTTTCTCACCTCGCGGCATCAAAAGTACAGGACTGCAACTGCTCAATCT	421
Db	1207	AGCTGTTTCACTTTCTCGCTCGCGGCATGAGACAGTGCAGGACTGCAACTGCTCAATCT	1266
QY	422	ATCCCGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	476
Db	1267	ATCCCGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1321
RESULT 12			
AA	AAC86939	standard; DNA; 9595 BP.	
ID	XX		
AC	XX		
AC	AA	AAC86939;	
DT	DT	02-APR-2001 (first entry)	
DE	DE	Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 1b.	
XX	XX	Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;	
KW	KW	HCV; vaccine; viral inhibitor; antiviral; ss.	
XX	XX	Hepatitis C virus.	
XX	XX		
FH	Key	Location/Qualifiers	
FT	CDS	342..9374	
FT		/*tag= a	
XX	XX	WO200075352-A2.	
XX	XX	14-DEC-2000.	
PD	PD	02-JUN-2000; 2000WO-US15527.	
XX	XX	04-JUN-1999; 99US-0137817.	
XX	XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PI	PI	Nam J, Bukh J, Emerson SU, Purcell RH;	
DR	DR	WPI; 2001-071081/08.	
DR	DR	P-PSDB; AAB31170.	
XX	XX	New nucleic acid comprising a chimeric bovine viral diarrhoea virus	
PT	PT	genome in which the (non-)structural region has been replaced by	
PT	PT	hepatitis C virus (HCV) genome useful for treating or preventing HCV	
XX	XX	signs and symptoms	
PS	PS	Disclosure; Fig 4A-F; 97pp; English.	
XX	XX	The specification describes a nucleic acid comprising a chimeric virus	
CC	CC	genome, specifically bovine viral diarrhoea virus (BVDV) genome in which	
CC	CC	the (non-)structural region has been replaced by the (non-)structural	
CC	CC	region of a hepatitis C virus (HCV) genome. The nucleic acids comprising	
CC	CC	the chimeric virus and the chimeric virus are useful for identifying	
CC	CC	cell lines capable of supporting the replication of these chimeric	
CC	CC	viruses, in screening for neutralizing antibodies to HCV of different	
CC	CC	genotypes, in the production of HCV-BVDV viruses, for the development	
CC	CC	of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal,	
CC	CC		

in studying the molecular properties of HCV indirectly in vitro, and in identifying inhibitors of viral enzyme activity which would be useful as antiviral agents. Formulations or compositions comprising the chimeric virions may be used to treat or prevent the signs and symptoms of HCV. The present sequence represents a HCV clone, which is used to construct chimeric nucleic acids of the invention.

Qy	302	CTGTTTCTGTCGGCTATGTACGTGGGGATCTCTGGGATCTGTGTTTCTGTTTCCC	361
Db	362	CTGCTTCTGCTCCGCTATGTACGTGGGAGATCTCTGGGATCTGTGTTTCTGTTTCCC	421
Qy	362	AGCTGTTCACTTCTCACCTTCGCCGCATCAAACAGTACAGGACTGCAACTGCTCAATCT	421
Db	422	AGCTGTTCACTTCTCACCTTCGCCGCATGAGACAGTACAGGACTGCAATTGCTCAATCT	481
Qy	422	ATCCGGCCATGTATCAGGTACCGCATGGCTTGGGATATGATGATGAATGGTTC	476
Db	482	ATCCTGGCCACGTATCAGGTATCGCATGGCTTGGGATATGATGATGAATGGTTC	536

Search completed: December 19, 2003, 18:51:05
Job time : 135.529 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1242.55 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-9

Perfect score: 483

Sequence: 1 ATGCCGGTGTCTTCTC.....TGATGACTGGTCCTAATAG 483

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estinu:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hcc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67.2	13.9	488	9 AV755731	AV755731 AV755731
C 2	55.6	11.5	492	9 AV758366	AV758366 AV758366
C 3	43.2	8.9	534	14 CD040840	CD040840 p9HB036XB
C 4	41.6	8.6	664	29 B2645446	B2645446 OGCBJ86TC

C	5	40.6	8.4	526	9	AL825643
	6	40.2	8.3	399	9	AV638521
	7	40.2	8.3	434	9	AV637507
	8	40.2	8.3	440	9	AV637983
	9	40.2	8.3	450	9	AV637259
	10	40.2	8.3	451	9	AV637328
	11	40.2	8.3	451	9	AV637643
	12	40.2	8.3	453	9	AV634724
	13	40.2	8.3	454	9	AV637050
	14	40.2	8.3	456	9	AV635382
	15	40.2	8.3	473	9	AV632765
	16	40.2	8.3	481	9	AV635503
	17	40.2	8.3	485	9	AV632811
	18	40.2	8.3	506	9	AV392445
	19	40.2	8.3	508	9	AV634095
	20	40.2	8.3	526	9	AV641895
	21	40.2	8.3	533	9	AV638125
	22	40.2	8.3	537	9	AV632335
	23	40.2	8.3	588	9	AV387329
	24	40	8.3	1186	13	BX421743
	25	39.4	8.2	1201	9	AL565958
C	26	39.2	8.1	624	14	CD206870
C	27	39.2	8.1	656	14	CB924688
	28	39	8.1	497	9	AV633658
	29	39	8.1	610	14	CB657655
	30	39	8.1	856	29	BZ578381
	31	39	8.1	872	29	BZ555011
C	32	38.4	8.0	645	29	CNS01213
C	33	38	7.9	771	29	BZ530934
	34	37.8	7.8	490	9	AV634529
C	35	37.8	7.8	705	14	CA618797
	36	37.8	7.8	1039	13	BX415186
C	37	37.6	7.8	309	12	BI098866
C	38	37.6	7.8	431	9	AV639153
C	39	37.4	7.7	394	12	BJ209795
	40	37.4	7.7	501	9	AV638474
C	41	37.2	7.7	431	9	AV641448
C	42	37.2	7.7	792	13	BX391120
	43	37	7.7	910	29	CNS0060N
	44	36.8	7.6	214	12	BM686105
	45	36.8	7.6	913	14	CA487901

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 BM Homo sapiens cDNA clone BMFAK03 5', mRNA linear EST 19-OCT-2000
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
TITLE Homo sapiens cDNA BM clones
JOURNAL Unpublished
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES

Mon Dec 22 13:28:55 2003

```
source 1..488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN
Query Match 13.9%; Score 67.2; DB 9; Length 488;
Best Local Similarity 67.0%; Pred. No. 1.1e-07;
Matches 126; Conservative 0; Mismatches 58; Indels 4; Gaps 2;

QY 292 GTTGGGGCTGCTTCTCTGTTCTGCTATGCTGGGGATCTCTGGGATCTCTGGGATCTGTTTC 351
DB 472 GTGGTGCACACTCGCTCTGCTCAGCTCTCTAGCTGGGACCTCTGGCAGGAGTGG 413
QY 352 GTTGTTCACAGCTGTTCACTTCTACCTCGCGGCATCAACAGTACAGGACTGCAAC 411
DB 412 CTTCAGTTCAGTGATCA---TCTGGCCTCAGCACCAGTGTGTCATGATGATCAAC 356
QY 412 TGCTCATCTATCCGGCCATGTATCAGTCAACGCTATG---GCTTGGGATATGATGAA 470
DB 355 TGCTCATCTATCTGGGGCCATCACTGGACCGTATGACGATGGGACATGATGATGAA 296
QY 471 CTGGTCTCT 478
DB 295 CTGGTCTG 288

RESULT 2
AV758366/c 492 bp mRNA linear EST 19-OCT-2000
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAKA03 5', mRNA sequence.
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng
, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1..492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAKA03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN

Query Match 8.9%; Score 43.2; DB 14; Length 534;
Best Local Similarity 47.4%; Pred. No. 0.39; 143; Indels 0; Gaps 0;
Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 70 TATGAAGTGCACAGTCTCGGGGTGTACCATCTACGACGACTGCTGCTCAACTCAAGC 129
DB 200 TACGGCGTGGCGAGATTACGGTATCGCATGGCTTCGCCGCTTCTACAACTGGACC 259
QY 130 ATAGTGTATGAGGCGAGCATGATCATGCACACCCCGGGTGGTGGCTTGGCTTGG 189
```



```

Db      260 TCATGACCAAGAGAGCCCGATCATGCTGACCCCAAGACGGTGCCCAACATTAC 319
QY      190 GAGGGCAACTCTCCGTTGCTGGGTGGCGCTCACTCCACAGCTCGCGCCAGGAAAGCC 249
Db      320 CACTAGCGGGACCACTCTCGGCTCGACCGTGGTGGCTTCGACGTGACAGATATATC 379
QY      250 AGGTGCCCCAACAGCAATAGACAGCCAGCTCGATTTGCTGTTGGGGTCTGCTGTTTC 309
Db      380 AACTTCTGACGACGAACCGGCTCTCGCAGGTATAGTGATCGGCGGTGACGCGACCCAC 439
QY      310 TGTTCGCTATGTACGTGGGGGATCTCTGGG 341
Db      440 CGTGCCGCCAACAGATCTCGGAGAGTGCCG 471

RESULT 4
LOCUS      BZ645446/c
DEFINITION BZ645446
ACCESSION BZ645446
VERSION   BZ645446.1
KEYWORDS  GSS
SOURCE    Zea mays
ORGANISM  Zea mays
REFERENCE 1 (bases 1 to 664)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
           A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
           R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE     Consortium for Maize Genomics
JOURNAL   Unpublished
COMMENT   Other GSSs: OGCBJ86TM
           Contact: Cathy Whitelaw
TIGR      9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES   Location/Qualifiers
            1..664
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMBMa0133P04"
            /clone_lib="ZM 0.7 1.5 KB"
            /notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"
BASE COUNT 103 a 238 c 222 g 101 t
ORIGIN

Query Match      8.6%; Score 41.6; DB 29; Length 664;
Best Local Similarity 50.5%; Pred. No. 1.1;
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY      119 CCAACTCAAGCATAGTGTATGAGCGGACGACATGATCATGCACACCCCGGGTGGCTGC 178
Db      266 CCGACGACGACACACGTGGTGGCGCCCAAGACGAGGTGCGGACGAGGCGCGGAC 207
QY      179 CTGCGGTTGGGAGGGCACTCTCCGTTGCTGGGTGGCGCTCACTCCACAGCTCGCGG 238
Db      206 CGTACCTGCGGTGGCGAGCTCCACGAAGTGGTAGAAGCCCTGCGCGTCTGTTGCGG 147
QY      239 CAGGAACCGCAGGTCCTCCACACGACATACGACGCCACGCTGATTCTGCTGTTGGG 298
Db      146 ACAGGAAGGCGCAGCGTCTCCACGACGACACCCACGAGCGGTACCCCTTCTCGGTGGT 87

```

```

QY      299 CTGCTGCTTCTGTTCCGCT 318
Db      86 GTGGTGGTGGTGGTGGTGGT 67

RESULT 5
LOCUS      AL825643/c
DEFINITION AL825643 p:234 Triticum aestivum cDNA clone A09_p234_plate_14, mRNA
           sequence.
ACCESSION AL825643
VERSION   AL825643.1
KEYWORDS  EST
SOURCE    Triticum aestivum (bread wheat)
ORGANISM  Triticum aestivum
REFERENCE 1 (bases 1 to 526)
AUTHORS   Wilson,I., Bewick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,
           Edwards,D., Coghill,J., Holdsworth,M., Lenton,J., Shewry,P. and
           Edwards,K.
TITLE     A BBSRC-funded wheat EST resource for the academic community
JOURNAL   Unpublished
COMMENT   Contact: Barker G
           Institute of Arable Crop Research
           Long Ashton, Bristol BS41 9AF United Kingdom.
FEATURES   Location/Qualifiers
            1..526
            /organism="Triticum aestivum"
            /mol_type="mRNA"
            /cultivar="mercia"
            /db_xref="taxon:4565"
            /clone="A09_p234_plate_14"
            /tissue_type="drought stressed seedlings"
            /dev_stage="21 days old"
            /clone_lib="p:234"
BASE COUNT 97 a 179 c 164 g 86 t
ORIGIN

Query Match      8.4%; Score 40.6; DB 9; Length 526;
Best Local Similarity 57.5%; Pred. No. 2;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY      150 CATGATCATGCACACCCCGGGTGGTGGCTCGCTCGTTCGGGAGGCAACTCTCTCCCGTTG 209
Db      151 CCTCTCCCGAACCCCGGGTGGATCAGTCGGAGTTCTTCGGCAGCCCTCTCTGCC 92
QY      210 CTGGGTGGGCTCACTCCACGCTCGCGCCAGGAACGCGAGCTCCCAACAGCAAT 269
Db      91 CTGGCGGGCGGCGGCGGCGGCGGTCGCGACAGAAACACGACGCGCGCGGCGGAC 32
QY      270 ACGACGC 276
Db      31 CACACCC 25

RESULT 6
LOCUS      AV638521
DEFINITION AV638521 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
           cDNA clone HC087d07_r 5', mRNA sequence.
ACCESSION AV638521
VERSION   AV638521.1
KEYWORDS  EST
SOURCE    Chlamydomonas reinhardtii
ORGANISM  Chlamydomonas reinhardtii
REFERENCE 1 (bases 1 to 399)
AUTHORS   Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
           Nakamura,Y. and Tabata,S.
TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2

```


JOURNAL adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
PUBMED 11089912

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 252-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers

FEATURES source
1..399
 /organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="C9"
 /db_xref="taxon:3055"
 /clones="HC087d07_r"
 /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"

BASE COUNT 70 a 144 c 122 g 63 t

ORIGIN
Query Match 8.3%; Score 40.2; DB 9; Length 399;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGTGTCCTGCTGCATACACAGCTTCCGCTTATGAAGTCGCAACGTGTC 89
 |||||
DB 44 CTTGCCACCCTCGGACGGCTCGTCATCGTGCGACCAACTTCGCCGTGGCACCATCTT 103
 |||||
QY 90 CGSGGGTAGCATGTGTCAGAACGACTGCTCCAACCTAAGCATAGTGTATGAGCGACGGGA 149
DB 104 CGCGGTGTGGCCCGGGGACAGCTGACCAACATCACCGCGCGAGCAGTGGCTGCGCG 163

QY 150 CATGATCATGCACAACCCCGGTGCGTCCCTCGGTTTCGGAGGGCAACTCTCCCGGTG 209
DB 164 CATGGGCATCTACGGTCCCACCGTGTCTTGTCAITTGCCCTGAAGACGCCCCGGCTG 223

QY 210 C 210
DB 224 C 224

RESULT 7
AV637507 LOCUS AV637507 Chlamydomonas reinhardtii 5' CO2 mRNA linear EST 15-DEC-2000
DEFINITION cDNA clone HC074a01.r 5', mRNA sequence.
ACCESSION AV637507
VERSION AV637507.1 GI:10780827
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales;
1 (bases 1 to 434)
Asamura,E., Miura,K., Kucho.K., Inoue,Y., Fukuzawa,H., Ohya,K., Nakamura.Y. and Tabata,S.

REFERENCE Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
AUTHORS DNA Res. 7 (5), 305-307 (2000)

TITLE

JOURNAL 20539644
MEDLINE 11089912
PUBMED

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/
Location/Qualifiers

FEATURES source
1..434
 /organism="Chlamydomonas reinhardtii"


```

/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC068a04"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/notes="vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon

```

BASE COUNT	82 a	153 c	144 g	75 t	
ORIGIN					
Query Match	8.3%	Score 40.2;	DB 9;	Length 454;	
Best Local Similarity	51.4%;	Pred. No. 2.4;			
Matches 93;	Conservative 0;	Mismatches 88;	Indels 0;	Gaps 0;	
Qy	30	CTTGGCCCTGCTGTCGTGACCATACAGACTTCGGCTTATGAAGTCGCAACGTGTC	89		
Db	52	CTTCGACCCCTGGACGGCTCGTCCATCGTCGACACCAACTTCGCCGTGGGCACCATCTT	111		
Qy	90	CGGGGTGATACCATGTCACGACAGACTCTCTCCAACCTCAAGCATAGTGTATGAGCAGCGGA	149		
Db	112	CGGGGTGTGCCCGGCGCAAGCTGACCAACATCACCGGCGGAGCAGGTGGCTGCCGG	171		
Qy	150	CATGATCATGCACACCCCGGGTGCGTGCCTCGGTTTCGGAGGGCAACTCTCCCGCTTG	209		
Db	172	CATGGGCATCTACGGTCCCCGCACCGGTGTTCTGCATTGCCCTGNAGACGCCCGCGCTG	231		
Qy	210	C 210			
Db	232	C 232			

RESULT 14	
AV635382	
LOCUS	456 bp mRNA linear EST 15-DEC-2000
DEFINITION	AV635382.Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii cDNA clone HC045f10_r 5', mRNA sequence.
ACCESSION	AV635382
VERSION	AV635382.1 GI:10778702
KEYWORDS	EST.
SOURCE	Chlamydomonas reinhardtii
ORGANISM	Chlamydomonas reinhardtii
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.
AUTHORS	1 (bases 1 to 456) Asamizu E., Miura K., Kucho K., Inoue Y., Fukuzawa H., Ohyama K., Nakamura Y. and Tabata S.
TITLE	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
JOURNAL	DNA Res. 7 (5), 305-307 (2000)
MEDLINE	20539644
PUBMED	11089912
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/ . Location/Qualifiers
FEATURES	

```

1. 436
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC045f10_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="vector: pBluescriptII SK-; Site1: EcoRI; Site2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
75 a      158 c      148 g      75 t

BASE COUNT
ORIGIN

```

Query Match	8.3%;	Score 40.2;	DB 9;	Length 456;
Best Local Similarity	51.4%;	Pred. No. 2.4;		
Matches 93;	Conservative 0;	Mismatches 88;	Indels 0;	Gaps 0;
Qy	30	CTTGGCCCTGCTGTCTGTCTGACCATACAGCTTCGGCTTATGAAGTGGCGAACGTC	89	
Db	221	CTTCGACCCCTGGACGGCTCGTCCATCGTCGACACCAACTTCGCCGTGGCCACCATCTT	280	
Qy	90	CGGGGTGATCATGTCTACGAACGACTGCTCCAACTCAAGCATATGATGTATGAGGACGGGA	149	
Db	281	CGGCGTGTGGCCGGGCACAAAGCTGACCAACATCACGGCCGGGAGCAGGTGGCTGCCGG	340	
Qy	150	CATGATCATGCACACCCCGGGTGGTGGCCCTCGGTTTCGGGAGGGCAACTCTCCCGGTG	209	
Db	341	CATGGGCATCTACGGTCCCGCCACCGTGTTCTGCATTGCCCTGAAGACGCCCGCGCTG	400	
Qy	210	C 210		
Db	401	C 401		

RESULT	15
AV632765	
LOCUS	AV632765 linear EST 15-DEC-2000
DEFINITION	Chlamydomonas reinhardtii 5% CO ₂ Chlamydomonas reinhardtii cDNA clone HC012c10_r 5', mRNA sequence.
ACCESSION	AV632765
VERSION	AV632765.1 GI:10776085
KEYWORDS	EST.
SOURCE	Chlamydomonas reinhardtii
ORGANISM	Chlamydomonas reinhardtii Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas. 1 (bases 1 to 473) Asamizu,Y., Miura,K., Kucho,K., Inoue.Y., Fukuzawa,H., Ohyama.K., Nakamura,E. and Tabata,S. Generation of expressed sequence tags from low-CO ₂ and high-CO ₂ adapted cells of Chlamydomonas reinhardtii
TITLE	DNA Res. 7 (5), 305-307 (2000)
JOURNAL	20539644
MEDLINE	11089912
PUBMED	Contact: Erika Asamizu The first Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/ .
COMMENT	Location/Qualifiers 1..473
FEATURES	source

```

0: .473
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC012c10_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/notes="Vector: pBluescriptII SK-; Site_- EcoRI; Site_2: xhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"
82 a 160 c 151 g :80 t

BASE COUNT
ORIGIN

Query Match      8.3%; Score 40.2; DB 9; Length 473;
Best Local Similarity 51.4%; Pred.No.2.5;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 30 CTTGGCCCTGTCGTCTGTGACCATACCAGCTTCGGCTTATGAAGTCGCACAAGTGTC 89
Db 156 CTTGCAGCCCTGGACCGCTGTCTCATCTGTGCACACCAACTTCGCCGTGGCACCATCTT 215

Qy :90 CGGGGTGTACATGTTCACGAACGACTGCTCCAATCAAGCATATGTGTAGGCGACGGGA 149
Db 216 CGGGCTGTGCGCGGCACAAAGCTTGACCAACATCACCGCGCGGACAGGTGGCTGCGGG 275

```


Mon Dec 22 13:28:55 2003

```
QY      150 CATGATCATGCACACCCCGGGTGGTGCCCTGCGTTGGGAGGGCAACTCCTCCCGTTG 209
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      276 CATGGGCATCTACGGTCCCGGCACCCGTGTTCTGCAATTGCCCTGAAGGACGCCCCCGGCTG 335
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      210 C 210
      |
Db      336 C 336
      |
```

Search completed: December 20, 2003, 06:54:43
Job time : 1246.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 34.8338 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-9

Perfect score: 483

Sequence: 1 ATCCCGGTGCTCTTCTC.....TGATGAACGTGCTCTAATAG 483

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	483	100.0	483	3	US-08-612-973-9
2	483	100.0	483	3	US-08-927-597-9
3	474.2	98.2	480	3	US-08-612-973-11
4	474.2	98.2	480	3	US-08-927-597-11
5	433.4	89.7	1539	2	US-08-470-426B-17
6	433.4	89.7	1863	2	US-08-470-426B-14
7	430.2	89.1	9595	3	US-09-014-416-4
8	430.2	89.1	9599	3	US-09-014-416-6
9	425.4	88.1	9472	4	US-08-150-204E-96
10	425	88.0	501	2	US-08-483-695-30
11	425	88.0	501	2	US-07-965-285-30
12	425	88.0	501	2	US-08-487-231-30
13	425	88.0	501	3	US-09-201-912-30
14	424.8	88.0	642	3	US-08-612-973-3
15	424.8	88.0	642	3	US-08-927-597-3
16	423.8	87.7	795	3	US-08-612-973-5
17	423.8	87.7	795	3	US-08-927-597-5
18	423.8	87.7	2082	3	US-08-612-973-47
19	423.8	87.7	2082	3	US-08-927-597-47
20	423.8	87.7	2433	3	US-08-612-973-49
21	423.8	87.7	2433	3	US-08-927-597-49
22	423.6	87.7	633	3	US-08-612-973-7
23	423.6	87.7	633	3	US-08-927-597-7
24	422.8	87.5	636	3	US-08-612-973-13
25	422.8	87.5	636	3	US-08-927-597-13
26	422.2	87.4	1037	1	US-08-462-195-1
27	422.2	87.4	1037	2	US-08-636-883-1

28 422.2 87.4 1037 3 US-09-127-829-1 Sequence 1, Appli

29 419 86.7 742 1 US-08-081-072-18 Sequence 18, Appl

30 419 86.7 742 1 US-08-449-093A-18 Sequence 18, Appl

31 419 86.7 932 1 US-08-081-072-15 Sequence 15, Appl

32 419 86.7 932 1 US-08-449-093A-15 Sequence 15, Appl

33 417.4 86.4 2116 3 US-08-191-160-21 Sequence 21, Appl

34 415.4 86.0 501 2 US-08-483-695-28 Sequence 28, Appl

35 415.4 86.0 501 2 US-07-965-285-28 Sequence 28, Appl

36 415.4 86.0 501 3 US-08-487-231-28 Sequence 28, Appl

37 415.4 86.0 501 3 US-09-201-912-28 Sequence 28, Appl

38 407.8 84.4 1167 1 US-08-324-977-9 Sequence 9, Appli

39 407.8 84.4 1167 2 US-08-384-616-9 Sequence 9, Appli

40 407.8 84.4 1167 2 US-08-904-686A-9 Sequence 9, Appli

41 407.8 84.4 1499 1 US-09-315-850-9 Sequence 3, Appli

42 407.8 84.4 1499 1 US-08-324-977-3 Sequence 3, Appli

43 407.8 84.4 1499 2 US-08-384-616-3 Sequence 3, Appli

44 407.8 84.4 1499 2 US-08-904-686A-3 Sequence 3, Appli

45 407.8 84.4 1499 3 US-09-315-850-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-08-612-973-9

; Sequence 9, Application us/08612973

; Patent No. 6150134

; GENERAL INFORMATION:

; APPLICANT: MAERTENS, GEERT

; APPLICANT: BOSMAN, FONS

; APPLICANT: DE MARTYNOFF, GUY

; APPLICANT: BUYSE, MARIE-ANGE

; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/612.973

; FILING DATE: 11-MAR-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: BYRNE, THOMAS E.

; REGISTRATION NUMBER: 32,205

; REFERENCE/DOCKET NUMBER: 1487-10

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 483 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..480

; FEATURE:

; NAME/KEY: mat.peptide

; LOCATION: 1..477

APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 base pairs
TYPE: nucleic acid
TOPOLOGY: linear
STRANDEDNESS: single
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..477
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..474

US-08-612-973-11

Query Match 98.2%; Score 474.2; DB 3; Length 480;
Best Local Similarity 99.4%; Pred. No. 3.4e-124;
Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGCCCGGTTGCTCTTTCTCTATCTCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60
Db 1 ATGCCCGGTTGCTCTTTCTCTATCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60
Qy 61 GCTTCGCTTATGAAGTGCAGCGTGTCCGGGTGTACATGTACAGACGACTGTCTCC 120
Db 61 GCTTCGCTTATGAAGTGCAGCGTGTCCGGGTGTACATGTACAGACGACTGTCTCC 120
Qy 121 AACTCAAGCATAGTGTATGAGGAGCGAGCATGATGACACCCCGGGTGGTGGCC 180
Db 121 AACTCAAGCATAGTGTATGAGGAGCGAGCATGATGACACCCCGGGTGGTGGCC 180
Qy 181 TGGCTTCGGGAGGSCAACTCTCCCGTTGTGGTGGCGGTCACTCCACGCTCGCGGCC 240
Db 181 TGGCTTCGGGAGGSCAACTCTCCCGTTGTGGTGGCGGTCACTCCACGCTCGCGGCC 240
Qy 241 AGGACGCCAGCTCCCAACACACATACGACGCACTGCGATTTGCTGTGGGGCT 300
Db 241 AGGACGCCAGCTCCCAACACACATACGACGCACTGCGATTTGCTGTGGGGCT 300
Qy 301 GCTGCTTTCTTCCTCGGTATGCTAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360
Db 301 GCTGCTTTCTTCCTCGGTATGCTAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360

Qy 361 CAGCTGTTACCTTCTCACCTCCCGGCATCAACACAGTACAGGACTGCAACTGCTCAATC 420
Db 361 CAGCTGTTACCTTCTCACCTCCCGGCATCAACACAGTACAGGACTGCAACTGCTCAATC 420
Qy 421 TATCCCGGCCATGTATCAGGTCAACGATCGCTTGGGATATGATGATGAAGTGGTCTCA 479
Db 421 TATCCCGGCCATGTATCAGGTCAACGATCGCTTGGGATATGATGATGAAGTGGTCA 479

RESULT 4

US-08-927-597-11
Sequence 11, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..477
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..474

US-08-927-597-11
Query Match 98.2%; Score 474.2; DB 3; Length 480;
Best Local Similarity 99.4%; Pred. No. 3.4e-124;
Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGCCCGGTTGCTCTTTCTCTATCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60
Db 1 ATGCCCGGTTGCTCTTTCTCTATCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60
Qy 61 GCTTCGCTTATGAAGTGCAGCGTGTCCGGGTGTACATGTACAGACGACTGTCTCC 120
Db 61 GCTTCGCTTATGAAGTGCAGCGTGTCCGGGTGTACATGTACAGACGACTGTCTCC 120
Qy 121 AACTCAAGCATAGTGTATGAGGAGCGAGCATGATGACACCCCGGGTGGTGGCC 180
Db 121 AACTCAAGCATAGTGTATGAGGAGCGAGCATGATGACACCCCGGGTGGTGGCC 180
Qy 181 TGGCTTCGGGAGGSCAACTCTCCCGTTGTGGTGGCGGTCACTCCACGCTCGCGGCC 240
Db 181 TGGCTTCGGGAGGSCAACTCTCCCGTTGTGGTGGCGGTCACTCCACGCTCGCGGCC 240
Qy 241 AGGACGCCAGCTCCCAACACACATACGACGCACTGCGATTTGCTGTGGGGCT 300
Db 241 AGGACGCCAGCTCCCAACACACATACGACGCACTGCGATTTGCTGTGGGGCT 300
Qy 301 GCTGCTTTCTTCCTCGGTATGCTAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360
Db 301 GCTGCTTTCTTCCTCGGTATGCTAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360

Query Match 98.2%; Score 474.2; DB 3; Length 480;
Best Local Similarity 99.4%; Pred. No. 3.4e-124;
Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGCCCGGTTGCTCTTTCTCTATCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60
Db 1 ATGCCCGGTTGCTCTTTCTCTATCTCTTGGCCCTGCTGCTCTGTGACCATACCA 60
Qy 61 GCTTCGCTTATGAAGTGCAGCGTGTCCGGGTGTACATGTACAGACGACTGTCTCC 120

INFORMATION FOR SEQ ID NO: 96
SEQUENCE CHARACTERISTICS:
LENGTH: 9472 base pairs

QY	242	GGAA	CGCC	AGCG	TCC	CCCA	CAAC	GAACA	ATAG	CAG	CGCC	ACG	TCG	ATTT	GC	TTC	GT	TGG	GGC	CTG	301	
Db	277	GGNA	TGCG	ACG	CGT	CCC	CACT	ACG	ACA	ATAC	GAC	CGC	ACG	TCG	ATTT	GC	TTC	GT	TGG	GGC	CGG	336
QY	302	CTG	CTTT	TGTT	CCG	CTAT	GTAG	TG	GGGG	GGAT	CTCT	CGG	ATCT	GT	TTT	CC	TG	TTT	CCC	361		
Db	337	CTG	CTTT	TCT	GTCT	CGG	TAT	GTAG	TG	GGGG	GGAT	CTCT	CGG	ATCT	GT	TTT	CC	TG	TTT	CCC	396	
QY	362	AGC	TGTT	TCA	CTT	TCT	CA	CTCT	CGC	CGG	CACT	CAAA	CAG	TAC	AG	CACTG	CA	TGCT	CAAT	CT	421	
Db	397	AGT	GT	TTC	AC	TTT	CT	CG	CC	TCT	CG	CGG	CA	TG	AG	CACTG	CA	TGCT	CAAT	CT	456	
QY	422	AT	CCG	GCC	CA	TG	TAC	AG	TCA	CGG	CA	TG	CGT	TTC	GG	ATAT	GAT	GTA	GA	466		
Db	457	AT	CCG	GCC	CA	TTT	AT	CA	G	GT	TCA	CGG	CA	TG	CGT	TTC	GG	ATAT	GAT	GTA	501	

RESULT 13
US-09-201-912-30
; Sequence 30, Application US/09201912
; Patent No. 6210962
; GENERAL INFORMATION:
; APPLICANT: Brecht, Christian
; APPLICANT: Kremsdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnean, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner, Robert M.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,912
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,285
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05286-0001-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
DESCRIPTION: cDNA to genomic RNA
US-09-201-912-30

Query Match	88.0%	Score 425;	DB 3;	Length 501;
Best Local Similarity	94.6%	Pred. No. 2.3e-110;		
Matches 440;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;
QY	2	TGCCCGGTGCTTTCTCTATCTTCTCTTGGCCCTGCTGCTCTGCTGACCATACCAG	61	

```

RESULT 12
US-08-487-231-30
; Sequence 30, Application US/08487231
; Patent No. 5919454
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremadort, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Finnegan, Henderson, Farabow, Garrett &
;

```

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,231
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

PRIOR APPLICATION NUMBER: US 07/965,285
APPLICATION NUMBER: 07/965,285
FILING DATE: 18-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 06 882
FILING DATE: 06-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05286-0001-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
DESCRIPTION: CDNA to genomic RNA
US-08-487-231-30

Query Match      88.0%; Score 425; DB 2; Length 501;
Best Local Similarity 94.6%; Pred. No. 2.3e-110;
Matches 440; Conservative 0; Mismatches 25; Indels 0; Gaps 0

QY      2   TGCCCGGTGTCTTTCTATCTTCCTTCTGGCCCTGCTGTCTGTCTGTGACCATACAG 61
DB      37  TGCCCGGTGTCTTTCTATCTTCCTTCTGGCCCTGCTGTCTGTGACCATCCAG 96

QY      62  CTTCCGCTTATGAAGTGCACAGTGTCCGGGTGTACCATGTACGACGACTGCTCCA 121
DB      97  CTTCCGCTTATGAAGTGCACAGTGTCCGGGTATACCATGTACGACGACTGCTCCA 156

```

122	ACTCAAGCATAGTGATGATGAGGAGCGGACATGATCATGACACCCCGGGTGGTGCCCT	181
QY		
Db		
157	ACTCAAGCATGTGTTATGAGGAGCGGACATGATCATGCTACTCCGGGTGGTGCCCT	216
182	GGGTTGGGAGGSCAACTCCTCCGTTGTTGGGTGGCGCTCACTCCACCGCTGGGGCCA	241
QY		

Mon Dec 22 13:28:54 2003

```

;
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..639
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..636
;
; US-08-927-597-3
;
; Query Match 88.0%; Score 424.8; DB 3; Length 642;
; Best Local Similarity 93.3%; Pred. No. 2.8e-110; Indels 0; Gaps 0;
; Matches 444; Conservative 0; Mismatches 32;
;
; QY 1 ATGCCGGTTCCTTTCTCTATCTTCTCTCTGTGGCCCTGTCTCTCTGTCTGACCATCCA 60
; DB 1 ATGCCGGTTCCTTTCTCTCTATCTTCTCTGTGGCTTACTGTCTCTGTGACCATCCA 60
;
; QY 61 GCTTCGGCTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTGTCAGAACGACTGCTCC 120
; DB 61 GCTTCGGCTTATGAGTGGCAACGTGTCCGGATGTACCATGTGTCAGAACGACTGCTCC 120
;
; QY 121 AACTCAAGCATGTGTATGAGCAGCGGACATGATCATGACACCCCGGTGGGTGCC 180
; DB 121 AACTCAAGCATTTGTGTATGAGCAGCGGACATGATCATGACACCCCGGTGGGTGCC 180
;
; QY 181 TCGGTTCCGGAGGGCAACTCCTCCGGTTGCTGGGTGGCGCTCACTCCACGCTCGCGCC 240
; DB 181 TCGGTTCCGGAGGAAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGAGCT 240
;
; QY 241 AGGAACGCCAGCGTCCCAACACGACAATACAGACCGCAGTCGATTTGCTCGTTGGGGCT 300
; DB 241 AGGAACGCCAGCGTCCCAACACGACAATACAGACCGCAGTCGATTTGCTCGTTGGGGGG 300
;
; QY 301 GCTGCTTTCTGTTCGCTATGTACGTGGGGATCTCTGCGGATCTGTCTTCTGTTTCC 360
; DB 301 GCTGCTCTCTGTTCGCTATGTACGTGGGGATCTCTGCGGATCTGTCTTCTCGTCTCC 360
;
; QY 361 CAGCTGTTTCACTTCTTCACTCCCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATC 420
; DB 361 CAGCTGTTTCACTTCTCGCTCCCGCATAGACGCGTGGAGGACTGCAATTTGCTCAATC 420
;
; QY 421 TATCCGGGCATGTATCAGTCAACCGATGCTTTGGGATATGATGATGAATGCTGGTC 476
; DB 421 TATCCGGGCACATAACAGGTACCGTATGGCTTTGGGATATGATGATGACTGGTC 476

```

Search completed: December 20, 2003, 07:03:00
Job time : 35.8338 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 1817.41 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-11
Perfect score: 480
Sequence: 1 ATGTCGGTGTCTTTCTC.....TGATGATGAACGTGTAATAG 480

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sta.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sta.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_mus.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_man.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	480	100.0	480	6	A48673	A48673 Sequence 11
2	480	100.0	480	6	AR157328	AR157328 Sequence
3	480	100.0	480	6	AX452760	AX452760 Sequence
4	480	100.0	480	6	AX685012	AX685012 Sequence
5	474.2	98.8	483	6	A48671	A48671 Sequence 9
6	474.2	98.8	483	6	AR157327	AR157327 Sequence
7	474.2	98.8	483	6	AX452758	AX452758 Sequence
8	474.2	98.8	483	6	AX685010	AX685010 Sequence
9	436	90.8	9379	14	AF207766	AF207766 Hepatitis C
10	432.8	90.2	1880	14	HPCSTRJ4	D08832 Hepatitis C
11	432.8	90.2	2540	6	E04260	E04260 CDNA encodi
12	432.8	90.2	2540	6	E04805	E04805 CDNA to 5'
13	432.8	90.2	2540	6	E07391	E07391 cDNA encodi
14	432.8	90.2	9448	14	HPCJ483	D13558 Hepatitis C
15	431.2	89.8	1539	6	AR027786	AR027786 Sequence
16	431.2	89.8	1863	6	AR027783	AR027783 Sequence
17	430.8	89.8	9580	14	AF054250	AF054250 Hepatitis
18	429.6	89.5	8780	14	AF054257	AF054257 Hepatitis
19	429.6	89.5	8780	14	AF054258	AF054258 Hepatitis
20	428	89.2	8781	14	AF054255	AF054255 Hepatitis
21	428	89.2	8781	14	AF054254	AF054254 Hepatitis
22	428	89.2	9379	14	AF165052	AF165052 Hepatitis
23	428	89.2	9460	14	HPCJ491	D10750 Hepatitis C
24	428	89.2	9595	6	AR119832	AR119832 Sequence
25	428	89.2	9595	14	AF054247	AF054247 Hepatitis
26	428	89.2	9596	14	AF054249	AF054249 Hepatitis
27	428	89.2	9599	6	AR119833	AR119833 Sequence
28	426.4	88.8	577	6	E04085	E04085 gDNA encodi
29	426.4	88.8	618	14	HPCJ491	D28929 Hepatitis C
30	426.4	88.8	8779	14	AF054251	AF054251 Hepatitis
31	426.4	88.8	8780	14	AF054252	AF054252 Hepatitis
32	426.4	88.8	9410	14	HPCJ182	D50481 Hepatitis C
33	426.4	88.8	9595	14	AF054248	AF054248 Hepatitis
34	426.2	88.8	633	6	A48669	A48669 Sequence 7
35	426.2	88.8	633	6	AR157326	AR157326 Sequence
36	426.2	88.8	633	6	AX452756	AX452756 Sequence
37	426.2	88.8	633	6	AX685008	AX685008 Sequence
38	426	88.8	9377	14	AF207756	AF207756 Hepatitis
39	426	88.8	9435	14	AB049093	AB049093 Hepatitis
40	424.8	88.5	3296	14	AB008446	AB008446 Hepatitis
41	424.8	88.5	8780	14	AF054259	AF054259 Hepatitis
42	424.8	88.5	8781	14	AF054256	AF054256 Hepatitis
43	424.8	88.5	9361	14	AF483269	AF483269 Hepatitis
44	424.8	88.5	9369	14	AF165054	AF165054 Hepatitis
45	424.8	88.5	9379	14	AF165051	AF165051 Hepatitis

ALIGNMENTS

RESULT 1
A48673
LOCUS A48673 Sequence 11 from Patent WO9604385. linear PAT 07-MAR-1997
DEFINITION A48673
ACCESSION A48673
VERSION A48673.1 GI:2302386
KEYWORDS
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 480)
AUTHORS Maertens, G., Bosman, F., De, M.G. and Buyse, M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 11 15-FEB-1996;

BASE COUNT	85 a	149 c	123 g	123 t	
ORIGIN					
Query Match	100.0%	Score 480;	DB 6;	Length 480;	
Best Local Similarity	100.0%;	Pred. No. 5e-113;			
Matches 480;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGTCGGTTCCTTTCTCTATCTTCCTCTTGGCCCTGCTGCTGCTGTGACCATACCA	60		
Db	1	ATGTCGGTTCCTTTCTCTATCTTCCTCTTGGCCCTGCTGCTGCTGTGACCATACCA	60		
Qy	61	GCCTCCGCTTATGAAGTGGCGCAACGTGTCGGGGTGTACCATGTCAAGAACGACTGCTCC	120		
Db	61	GCCTCCGCTTATGAAGTGGCGCAACGTGTCGGGGTGTACCATGTCAAGAACGACTGCTCC	120		
Qy	121	AATCTAAGCATAGTGTATGAGGACGGGACATCATATGCACACCCCGGGTGGTGGCC	180		
Db	121	AATCTAAGCATAGTGTATGAGGACGGGACATCATATGCACACCCCGGGTGGTGGCC	180		
Qy	181	TGCGTTGGGAGGGCAACTCCTCCGTTGCTGGGTGGCGCTCACTCCACGCTGCGGGCT	240		
Db	181	TGCGTTGGGAGGGCAACTCCTCCGTTGCTGGGTGGCGCTCACTCCACGCTGCGGGCT	240		
Qy	241	AGGAAGCCAGCGTCCCAACACGACATACGACGACGCGCCACGTTCGTTGGGGCT	300		
Db	241	AGGAAGCCAGCGTCCCAACACGACATACGACGACGCGCCACGTTCGTTGGGGCT	300		
Qy	301	GCTGCTTTCTGTTCCCGTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTC	360		
Db	301	GCTGCTTTCTGTTCCCGTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTC	360		
Qy	361	CAGCTGTTCACCTTCTCACTCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATC	420		
Db	361	CAGCTGTTCACCTTCTCACTCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATC	420		
Qy	421	TATCCGGCCATGTATCAGTGTACCGCATCGCTTGGGATATGATGAACTGGTAATAG	480		
Db	421	TATCCGGCCATGTATCAGTGTACCGCATCGCTTGGGATATGATGAACTGGTAATAG	480		
RESULT 5					
A48671					
LOCUS	A48671	483 bp	DNA	linear	PAT 07-MAR-1997
DEFINITION	Sequence 9 from Patent WO9604385.				
ACCESSION	A48671				
VERSION	A48671.1	GI:2302384			
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 483)				
AUTHORS	Maertens, G., Bosman, F., De, M. G. and Buyse, M.				
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE				
JOURNAL	Patent: WO 9604385-A 9 15-FEB-1996;				
COMMENT	INNOCENTIS NV (BE)				
OTHER PUBLICATION	Other publication CA 2172273 960215				
OTHER PUBLICATION	Other publication AU 3382495 960304.				
FEATURES	Location/Qualifiers				
source	1..483				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
CDS	1..480				
	/code="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAA03131.1"				
	/db_xref="GI:2302385"				
	/translation="MPGCSFISLLALLSLCTIPASAYEVNRYGVHTVNDGSSSTI				
	YVRAADMIHTPGCPVCRGNSRRCWALTPILARNASVPTTIRRHVDLLVGAAL				
	FCSAMVGDVUCGSVFLVQFTFSPPRHQVTQDCNCSIIYPGVSHRMAWDMMNWIS				
	1..477				
	/product="unnamed"				
mat_peptide					

[illegible]

QY	422	ATCCCGCCATGTATCAGGTACCCGATGGCTTGGATATGATGATGAACCTGTTAA	477
Db	1255	ATCCCGCCACTTATCAGGTACCCGATGGCTTGGATATGATGATGAACCTGTTCA	1310
RESULT 10			
HPC5TRJ4			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
FEATURES			
source			
CDS			

QY	422	ATCCCGCCATGTATCAGGTACCCGATGGCTTGGATATGATGATGAACCTGTTAA	477
Db	1255	ATCCCGCCACTTATCAGGTACCCGATGGCTTGGATATGATGATGAACCTGTTCA	1310
RESULT 10			
HPC5TRJ4			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
FEATURES			
source			
CDS			

GVNATCNLPGCSFSLFLLALISCLTIPASAVEVRNVSGIYHVNDSCNSSIYVEAD
MIMHTPCVPCVREDNSSRCWALTPFLAARNASVETTTIRRHVDLLVGAAPCSAMY
VGLDGLSVFLVQLFETSPRHETVQDCNSIYPGHLSGHRMAWMMNMTTALV
SOLLRIPOAVDMVAGAHWGLAGLAYSMVGNWAKVLIVALLFAGVGETYSGAA
SHTTSLASLFPGRSQRQLVNTNGSMHNRNALCNDLSLHTGFLAALFYTHRFNS
GCPERMASCRPIDWFAQGWGPITYTEPDSPDQRPYCMHYAPRPGIVPASPQVCGPYC
FTPSP"

BASE COUNT 333 a 593 c 555 g 399 t

Query Match 90.2%; Score 432.8; DB 14; Length 1880;

Best Local Similarity 94.3%; Pred. No. 7.4e-101;

Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY	2	TGTCGGGTGCTTTCTCTATCTTCTTCTGCTTGTGCGCCCTGCTGCTGTCTGACCATACGAG 61
DB	847	TGCGCGGTGCTCTTTCTCTATCTTCTTCTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTG 906
QY	62	CTTCCGCTTATGAAGTGCACAGCTGTCGGGGGTGTACCATGTACGACAGCAAGACTGCTCCA 121
DB	907	CTTCCGCTTATGAAGTGCACAGCTGTCGGGGGTGTACCATGTACGACAGCAAGACTGCTCCA 966
QY	122	ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGACACACCCCGGGTGGTGGCCCT 181
DB	967	ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGATCTCCGGGTGGTGGCCCT 1026
QY	182	GGTTTCGGAGGCAACTCTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA 241
DB	1027	GGTTTCGGAGGCAACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA 1086
QY	242	GGAGCCGAGCGTCCCAACAGACATAGCAGCCAGCTGCAATTTGCTGTTGGGGCTG 301
DB	1087	GGAGCCGAGCGTCCCAACAGACATAGCAGCCAGCTGCAATTTGCTGTTGGGGCTG 1146
QY	302	CTGCTTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
DB	1147	CTGCTTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1206
QY	362	AGCTGTTCACCTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
DB	1207	AGCTGTTCACCTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 1266
QY	422	ATCCCGGCGCATGTATCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 477
DB	1267	ATCCCGGCGCATGTATCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1322

RESULT 11

E04260
LOCUS
DEFINITION
cDNA encoding a part of type non-A non-B hepatitis virus.
ACCESION
E04260
VERSION
E04260.1 GI:2172463
KEYWORDS
JP 1993023200-A/2.
SOURCE
unidentified
ORGANISM
unclassified.

REFERENCE
1 (bases 1 to 2540)
Okamoto, H. and Nakamura, T.

AUTHORS
Okamoto, H. and Nakamura, T.

TITLE
HIGHLY SENSITIVE DETECTION METHOD OF NON-A NON-B TYPE HEPATITIS

JOURNAL
VIRUS USING OLIGONUCLEOTIDE PRIMER AND OLIGONUCLEOTIDE PRIMER

COMMENT
Patent: JP 1993023200-A 2 02-FEB-1993;
NAKAMURA TETSUO

PN JP 1993023200-A/2

PD 02-FEB-1993

PF 26-APR-1991 JP 1991191376

PR 12-JUN-1990 JP 90P 153402

PI OKAMOTO HIROAKI, NAKAMURA TETSUO

PC C12Q1/68, C12N15/51, C12Q1/70;

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

*source: strain=HC-J4;
Key Location/Qualifiers

misc_feature 1..2540

/note='a part of type non-A non-B hepatitis

virus';

Location/Qualifiers

1..2540

/organism='unidentified'

/mol_type='genomic RNA'

/db_xref='taxon:32644'

BASE COUNT 470 a 776 c 741 g 553 t

ORIGIN

Query Match 90.2%; Score 432.8; DB 6; Length 2540;

Best Local Similarity 94.3%; Pred. No. 7.4e-101;

Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY	2	TGTCGGGTGCTTTCTCTATCTTCTTCTGCTTGTGCGCCCTGCTGCTGTCTGACCATACGAG 61
DB	847	TGCGCGGTGCTCTTTCTCTATCTTCTTCTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTG 906
QY	62	CTTCCGCTTATGAAGTGCACAGCTGTCGGGGGTGTACCATGTACGACAGCAAGACTGCTCCA 121
DB	907	CTTCCGCTTATGAAGTGCACAGCTGTCGGGGGTGTACCATGTACGACAGCAAGACTGCTCCA 966
QY	122	ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGACACACCCCGGGTGGTGGCCCT 181
DB	967	ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGATCTCCGGGTGGTGGCCCT 1026
QY	182	GGTTTCGGAGGCAACTCTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA 241
DB	1027	GGTTTCGGAGGCAACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA 1086
QY	242	GGAGCCGAGCGTCCCAACAGACATAGCAGCCAGCTGCAATTTGCTGTTGGGGCTG 301
DB	1087	GGAGCCGAGCGTCCCAACAGACATAGCAGCCAGCTGCAATTTGCTGTTGGGGCTG 1146
QY	302	CTGCTTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
DB	1147	CTGCTTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1206
QY	362	AGCTGTTCACCTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
DB	1207	AGCTGTTCACCTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 1266
QY	422	ATCCCGGCGCATGTATCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 477
DB	1267	ATCCCGGCGCATGTATCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1322

RESULT 12

E04805

LOCUS

DEFINITION

cDNA to 5'-terminal region of gRNA of Hepatitis nonA nonB virus.

ACCESION

E04805

VERSION

E04805.1 GI:2173001

KEYWORDS

JP 1993091884-A/2.

SOURCE

unidentified

ORGANISM

unclassified.

REFERENCE

1 (bases 1 to 2540)

Okamoto, H. and Nakamura, T.

AUTHORS

Okamoto, H. and Nakamura, T.

TITLE

DETECTION SYSTEM FOR NON-A NON-B HEPATITIS VIRUS RELATING ANTIGEN

JOURNAL

AND ANTIBODY, POLYNUCLEOTIDE AND POLYPEPTIDE

COMMENT

Patent: JP 1993091884-A 2 16-APR-1993;
NAKAMURA TETSUO

PN JP 1993091884-A/2

PD 16-APR-1993

PF 10-APR-1991 JP 1991196175

PR 12-JUN-1990 JP 90P 153401, 08-NOV-1990 JP 90P 304405 PI

OKAMOTO HIROAKI, NAKAMURA TETSUO

OS	Hepatitis non-A non-B virus
PN	JP 1994125777-A/4
PD	10-MAY-1994
PF	20-JUN-1991 JP 1991247120
PI	OKAMOTO HIROAKI, NAKAMURA TETSUO
PC	C12N15/51, C12N1/02, C12P21/02, C12Q1/70, G01N33/53, G01N33/569, PC G01N33/576;
CC	strandedness: Single;
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No;
PH	key
PH	Location/Qualifiers
FT	source
FT	1. .2540
FT	/organism='Hepatitis non-A non-B virus' FT
FT	/strain='HC-J4'
FT	5'UTR
FT	1. .341
FT	mat_peptide
FT	342..1490
FT	/note='non-structural protein of chimpanzee
FT	hepatitis non-A
FT	non-B virus'
FT	mat_peptide
FT	1491..2540
FT	/note='Structural protein of chimpanzee FT
FT	hepatitis non-A
FT	non-B virus'.
FEATURES	Location/Qualifiers
source	1. .2540
	/organism="unidentified"
	/mol_type="genomic RNA"
	/db_xref="taxon:32644"
BASE COUNT	470 a 776 c 741 g 553 t
ORIGIN	
Query Match	90.2%; Score 432.8; DB 6; Length 2540;
Best Local Similarity	94.3%; Pred. No. 7.4e-101;
Matches 449; Conservative	0; Mismatches 27; Indels 0; Gaps 0;
QY	2 TGTCGGGTGCTCTTCTCTATCTCTCTCTGGCCCTGCTGTCCTGTCTGACCATACCG 61
Db	847 TGCCCGGTGCTCTTCTCTAICTTCTCTTGGCTTGTCTGCTGTTGACCATCCCGAG 906
QY	62 CTTCGGCTTATGAAGTCGGCAACGTCGGGGTGTACCATGTCCAGACGACTGCTCCA 121
Db	907 CTTCGGCTTATGAAGTCGGCAACGTCGGGGATATACCATGTCCAGACGACTGCTCCA 966
QY	122 ACTCAAGCATGTATGAGCAGCGCATGATCATGCACACCCCGGGTGGTGCCTT 181
Db	967 ACTCAAGCATGTGTATGAGCAGCGCATGATCATGCATCTCCCGGGTGGTGCCTT 1026
QY	182 CGCTTCGGGAGGGCAACTCTCCCGTGTGTGGGTGGCGTCACTCCACGCTCGCGGCA 241
Db	1027 CGCTTCGGGAGGACAAACAGTCCCGTGTGTGGGTAGCGCTCACTCCACGCTCGCGGCA 1086
QY	242 GGAAGCCAGCGTCCCAACAGCAATACGACGCCACGTCGATTTGCTTGGGGCTG 301
Db	1087 GGAATGCCAGCGTCCCACTACGACAATACGCCACGTCGATTTGCTTGGGGCGG 1146
QY	302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 361
Db	1147 CTGCTTTCTGCTCCGCTATGTACGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 1206
QY	362 AGCTGTTCACCTTCTCAGCTCCGGGATCAACAGATACAGGACTGCACCTGCTCAATCT 421
Db	1207 AGCTGTTCACCTTCTCGCTCCGGGATGAGACGTGCAGGACTGCACCTGCTCAATCT 1266
QY	422 ATCCCGGCGCATGTATCAGGTACCGCATGGCTTGGGATGATGATGATGATGATGATG 477
Db	1267 ATCCCGGCGCATTTATCAGGTACCGCATGGCTTGGGATGATGATGATGATGATGATG 1322
RESULT 14	
HPCJ483	
LOCUS	
	9448 bp RNA linear VRL 01-FEB-2000

DEFINITION

Hepatitis C virus genome, complete sequence.

ACCESSION

D1358 D01217

VERSION

D1358.1 Gi:221604

KEYWORDS

C protein; E2 protein; NS1 protein; NS2 protein; NS3

SOURCE

Hepatitis C virus

ORGANISM

Hepatitis C virus

REFERENCE

1 (bases 1 to 9448)

AUTHORS

Okamoto, H., Kojima, M., Okada, S., Yoshizawa, H., Iizuka, H.,

Tanaka, T., Muchmore, E.E., Peterson, D.A., Ito, Y. and Mishiro, S.

TITLE

Genetic drift of hepatitis C virus during an 8.2-year infection in

a chimpanzee: variability and stability

JOURNAL

Virology 190 (2), 894-899 (1992)

MEDLINE

92391112

PUBMED

1325713

REFERENCE

2 (bases 1 to 9448)

AUTHORS

Okamoto, H.

TITLE

Direct Submission

JOURNAL

Submitted (17-Oct-1991)

COMMENT

Hiroaki Okamoto, Jichi Medical School,

Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi,

329-04, Japan (E-mail: hokamoto@jichi.ac.jp,

Tel: 0285-44-2111 (ex. 3334), Fax: 0285-44-1557)

Submitted (17-Oct-1991) to DDBJ by:

Hiroaki Okamoto

Immunology Division

Jichi Medical School

Kawachi-gun

Tochigi 329-04

Japan

Phone: 0285-44-2111 x3334

Fax: 0285-44-1557.

FEATURES

source

1. .9448

/organism="Hepatitis C virus"

/mol_type="genomic RNA"

/strain="HC-J4"

/isolates="HC-J4/83"

/db_xref="taxon:11103"

342_..9374

/codon_start=1

/product="polyprotein precursor"

/protein_id="BAA02756.1"

/db_xref="GI:221605"

/translation="MSTNPKPQKTKRNTNRPDQVKFPGGQIVGGVLLPRGPRL

GVRAETKTSRQPRGWRQPIPKARPEGRWAQPCYWPPLYGNEGLWAGWLLSPRG

SRPSGPTDPRRSRNLGVIDTLTCGFADMLGIIPLVGAAPLGRRAAHLGVRLLED

GVNATGNLPGCCSFSLIALLSCLTIPASAVEVRNVSGIYHVNDCSNISYVEAAD

MIMHTPGVPCVREDNSRCWALPTLAARNAVSPTTIRRHVDLLVGAAPFCSAMY

VGDLCGSLVSLQFLTFSPRRHETVQDCNCSYPGHLSGRHMAWMMWNSPTALVY

SOLLRTPOAVDMVAGAHVGLAGLAYYSMVGNAKVLIVALLFAGVDCETVTSGLAA

SHTSTLSLSPGASQRIQLVNTNGSWHINRTALNCNLSLHGTFLAALPYTHRFNS

GCPERNASCRPDWFAQGGPIITYEPDSPQRPICWRYAPRCGPVAPASQVGPVYC

FPTSPVNVGHTDSRGTYSWGENETVDMNNRPPQGNWFGCTWMNSTGFTKTCG

FPCNTGVGNHTLTCTPCFRKHEATYTKCGSPWLTPLRCLVDYPIRLWHYPTCFN

SIFKVMYVGVGHEIRLNAACNTRGRCLNDRSELSPLLSTTEWOLLPCAFTTL

PALSTGLIHLHNIQVQLYGVGSFVSFAIKWEYVILLFLLILADARVACILMMML

IAQAEPALVNLNAAGAHGILSLVFFCAWYIKGRAPAGAYFYGWVPLLL

LLALLPRAYMDREMAASCAGVAVGLVFLTLSPYIKVFLTLIWLQITFIIRAEAR

MQVWVPLNVGRGDRIILLTCAVHPELFDITKLKLLALILGLMLVQLQITRVFVR

AQLGRACMLVRAAGHYVQAFMKLGALTGTYYVYHNLTPLRDWAHAGRLDLAVE

PVVFSDMETKVTWGAQTAACGDIILGLPVSARRGKEIFLGPADSLEGOGWRLLAIT

AYSQOVRGLVGLCIITSLTRGDKNOVEGVQVVSATQSFATLCINGVCTVTVHAGSK

TLAGPKGPIQMYTNVDLILVGNQAPPGARSMTPTCGSSDLVLVTRHADVIVRRRG

DSRGLSPRPVYLKSGSGGLLCPSPGVGVVGFRAAVCTRGVAKADFIPIVSMETT

MRSPVTDNRSPVQPTQVAHLHAPTSGSKTKVPAAYAGQYKVLVNFPSVAATL

GFGAYMPKAGHDIPNRTGVTITITGGPIITYSTYKFLADGGSGGAYDIIICDECHS

TOSTTILGTLGTDOETAGARLVLAATATPGSVTVPHNIBIEIGLSNNGEIPFYGK

AIPLEAKGSRHILFCHSKKCDLAKLTGLGLNAVAYYRGLDVSVIPIIGDAVVA

TDALMTGFGDFSDVDCNTCVQTVDFSLDPTFTTETTTVPDASVRSQRRGTORG

RSGIYRFTVPGSRPSGMFSSVLCEDYDAGCAWYELTPAETSVRLRAYLNTPLGPVQC

DHLFEVESVFTGLSHIDAHFLSQTKQAGDNFPYLVAYQATVCARAAPPSPDWQWKC

mat_peptide 342..914

/product="C protein"

mat_peptide 915..1490

/product="E protein"

mat_peptide 1491..2528

/product="NS1/E2 protein"

mat_peptide 2529..3359

/product="NS2 protein"

mat_peptide 3360..5186

/product="NS3 protein"

mat_peptide 5187..6380

/product="NS4 protein"

mat_peptide 6381..9371

/product="NS5 protein"

stem_loop 6005..6104

/notes="possible stem-loop structure in NS4 sequence"

misc_feature 9413..9448

/note="T-stretch of 36 bp"

BASE COUNT 1899 a 2828 c 2676 g 2045 t

ORIGIN

Query Match 90.2%; Score 432.8; DB 14; Length 9448;

Best Local Similarity 94.3%; Pred. No. 7.6e-101;

Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

2 TGTCCGGTTGCTCTTTCTCTATCTTCTTGGCCCTGCTGCTCTGACCATACCAG 61

847 TGGCCGGTTGCTCTTTCTCTATCTTCTTGGCTTGTCTGCTTGTGACCATCCAG 906

62 CTTTCCGCTTATGAAGTGCACACAGTGTCCGGGGTGATACCATGTCAAGAACAGTGTCCA 121

907 CTTCCGCTTATGAAGTGCACACAGTGTCCGGGATATACCATGTCAAGAACAGTGTCCA 966

122 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGTGCGCCT 181

967 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCATATCTCCCGGGTGTGCGCCT 1026

182 GCCTTCGGAGGCGCAACTCTCCCGTGTGGGTGGCGCTCACTCCACAGCTCCGCGCCA 241

1027 GCCTTCGGAGGCGCAACTCTCCCGTGTGGGTGGCGCTCACTCCACAGCTCCGCGCCA 1086

242 GGAACGCCAGCGTCCCAACACAAATACGACGCCACGTCGATTTGCTGTTGGGGCTG 301

1087 GGAATGCCAGCGTCCCAACACAAATACGACGCCACGTCGATTTGCTGTTGGGGCTG 1146

302 CTGCTTTCTGTTCCGCTATGACGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 361

1147 CTGCTTTCTGTTCCGCTATGACGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 1206

Mon Dec 22 13:28:32 2003

```

QY      362 AGCTGTTACCTTCTCCTCGCCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db      1207 AGCTGTTACCTTCTCGCCCTCGCCGGCATGAGACAGTGCAGGACTGCAACTGCTCAATCT 1266

QY      422 ATCCCGGCCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGAAGTGGTAA 477
Db      1267 ATCCCGGCCCATTTATCAGGTACCGCATGCTTGGGATATGATGATGAAGTGGTCA 1322

RESULT 15
AR027786      AR027786      1539 bp      DNA      linear      PAT 29-SEP-1999
LOCUS      Sequence 17 from patent US 5856458.
DEFINITION      AR027786
ACCESSION      AR027786
VERSION      AR027786.1      GI:5938606
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      Unclassified.
AUTHORS      1. (bases 1 to 1539)
TITLES      Okamoto,H. and Nakamura,T.
JOURNAL      Oligonucleotide primers, and their application for high-fidelity
FEATURES      detection of non-A, non-B hepatitis virus
              Patent: US 5856458-A 17 05-JAN-1999;
              Location/Qualifiers
              source      1..1539
              /organism="unknown"
BASE COUNT      271 a 490 c 448 g 330 t
ORIGIN

Query Match      89.8%; Score 431.2; DB 6; Length 1539;
Best Local Similarity 94.1%; Pred. No. 1.9e-100;
Matches 448; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY      2 TGTCGGGTGCTCTTTCTCTATCTCTCTTGGCCCTGCTGCTGTGACCATACCAG 61
Db      506 TGCCCGGTGCTCTTTCTCTATCTCTCTTGGCTTTGCTGTCTGTGACCATCCAG 565

QY      62 CTTCGGCTTATGAAGTGGCAACGTGTCGGGGTGTACCATGTCAAGACGACTGTCCA 121
Db      566 CTTCGGCTTATGAAGTGGCAACGTGTCGGGATATACCATGTCAAGACGACTGTCCA 625

QY      122 ACTCAAGCATGTGTATGAGGAGCGGACATGATCATGACACACCCCGGGTGGTGCCT 181
Db      626 ACTCAAGCATTTGTATGAGGAGCGGACATGATCATGACTACTCCCGGGTGGTGCCT 685

QY      182 GCGTTGGGAGGCAACTCCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGCCA 241
Db      686 GCGTTGGGAGGCAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCACGCTCGCGCCA 745

QY      242 GGAACGCCACGCTCCGCCACAACAGCAATAACGACGACGACGACGACGACGACGACG 301
Db      746 GGAATGCCAGGGTCCCGCTACGACAAATACGACGACGACGACGACGACGACGACGACG 805

QY      302 CTGCTTTCTGTCGGGTATGTAGTGGGGATCTCTGGGATCTGTTTCTGTTTCCC 361
Db      806 CTGCTTTCTGTCGGGTATGTAGTGGGGATCTCTGGGATCTGTTTCTGTTTCCC 865

QY      362 AGCTGTTACCTTCTCACCCTCGCCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db      866 AGCTGTTACCTTCTCGCCCTCGCCGGCATGAGACAGTGCAGGACTGCAACTGCTCAATCT 925

QY      422 ATCCCGGCCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGAAGTGGTAA 477
Db      926 ATCCCGGCCCATTTATCAGGTACCGCATGCTTGGGATATGATGATGAAGTGGTCA 981
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 132.7 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-11

Perfect score: 480

Sequence: 1 ATGTCGGTGTCTTCTCTC.....TCGATGATGAACGTGTAATAG 480

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	17 AAT12708	HCV E1 construct H
2	480	100.0	480	24 AAL48917	Hepatitis C virus
3	474.2	98.8	483	17 AAT12707	HCV E1 construct H
4	474.2	98.8	483	24 AAL48916	Hepatitis C virus
5	432.8	90.2	1880	13 AAQ24467	NANB hepatitis vir
6	432.8	90.2	2187	19 ABA03491	Cuticle protein 1
7	432.8	90.2	2540	14 AAQ43889	NANB hepatitis vir
8	432.8	90.2	2540	15 AAQ63753	NANSHV genomic fra

9	431.2	89.8	1863	12 AAQ15363	Fragment of NANB h
10	429.6	89.5	2540	13 AAQ29628	Hepatitis C virus
11	428	89.2	9595	20 AAX24843	Infectious hepatitis
12	428	89.2	9595	22 AAC86939	Nucleotide sequenc
13	428	89.2	9595	22 AAF23492	Infectious hepatitis
14	428	89.2	9599	20 AAX24833	Infectious hepatitis
15	426.4	88.8	577	14 AAQ35081	HCV envelope regio
16	426.4	88.8	2187	19 ABA03492	Cuticle protein 1
17	426.2	88.8	633	17 AAT12706	HCV E1 construct H
18	426.2	88.8	633	24 AAL48915	Hepatitis C virus
19	424.6	88.5	567	13 AAQ27160	NANB hepatitis vir
20	423.2	88.2	577	14 AAQ35085	HCV envelope regio
21	423.2	88.2	580	12 AAQ11076	Fragment of hepati
22	423.2	88.2	580	20 AAZ07647	HCV J1 E domain co
23	423.2	88.2	1562	19 AAV60672	Fragment #5 isolat
24	423.2	88.2	1953	25 AAL55222	Plasmid pIDKE2 DNA
25	423.2	88.2	2829	19 AAV60673	Fragment #6 isolat
26	422.8	88.1	580	16 AAQ79750	Hepatitis C virus
27	422.8	88.1	580	20 AAX26733	Consensus sequence
28	422.8	88.1	580	20 AAX26728	Consensus sequence
29	422.8	88.1	580	20 AAX00401	Hepatitis C virus
30	422.8	88.1	1249	16 AAQ79772	Hepatitis C virus
31	422.8	88.1	1249	20 AAX26739	Consensus sequence
32	422.8	88.1	1270	19 AAV60668	Fragment #1 isolat
33	422.8	88.1	3401	15 AAQ64069	Non-A, non-B hepat
34	422.8	88.1	3401	16 AAT30387	5'UTR/CORE/ENV/NS1
35	422.8	88.1	3461	15 AAQ64068	Non-A, non-B hepat
36	422.8	88.1	3461	16 AAT30386	5'UTR/CORE/ENV/NS1
37	422.2	88.0	642	17 AAT12704	HCV E1 construct H
38	422.2	88.0	642	24 AAL48913	Hepatitis C virus
39	421.6	87.8	1882	13 AAQ27159	NANB hepatitis vir
40	421.2	87.8	577	14 AAQ35086	HCV envelope regio
41	421.2	87.8	580	16 AAQ79758	Hepatitis C virus
42	421.2	87.8	580	16 AAQ79759	Hepatitis C virus
43	421.2	87.8	580	16 AAQ79760	Hepatitis C virus
44	421.2	87.8	795	17 AAT12705	HCV E1 construct H
45	421.2	87.8	795	24 AAL48914	Hepatitis C virus

ALIGNMENTS

RESULT 1
AAT12708
ID AAT12708 standard; DNA; 480 BP.

XX AAT12708;
XX
XX AAT12708;
XX
XX 23-SEP-1996 (first entry)
XX
XX HCV E1 construct HCC113A.
XX
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX ss.
XX
XX Hepatitis C virus.
XX
XX WO9604385-A2.
XX
XX 15-FEB-1996.
XX
XX 31-JUL-1995; 95WO-BF03031.
XX
XX 29-JUL-1994; 94EP-0870132.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
XX WPI; 1996-129401/13.
XX
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope


```

Db 361 GCTGCTTACCTTTCACCTGCGGCATCAACAGTAGTACAGACTGCAACTGCTCAATC 420
Qy 421 TATCCCGGCATGATCAGCTACCGCATGGCTTGGGATATGATGAACTGGTAATAG 480
Db 421 TATCCCGGCATGATCAGCTACCGCATGGCTTGGGATATGATGAACTGGTAATAG 480

RESULT 3
AAT12707
ID AAT12707 standard; DNA; 483 BP.
XX AC AAT12707;
XX DT 23-SEP-1996 (first entry)
XX DE HCV E1 construct HCC112A.
XX KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX OS Hepatitis C virus.
XX PN WO9604385-A2.
XX PD 15-FEB-1996.
XX PF 31-JUL-1995; 95WO-EP03031.
XX PR 29-JUL-1994; 94EP-0870132.
XX PA (INNO-) INNOGENETICS NV.
XX PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX DR
XX PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di: sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX PS Claim 23; Fig 21; 146pp; English.
XX CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX SQ Sequence 483 BP; 85 A; 152 C; 123 G; 123 T; 0 other;

Query Match 98.8%; Score 474.2; DB 17; Length 483;
Best Local Similarity 99.4%; Pred. No. 9.4e-129;
Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGTCGGTTCCTTCTATCTATCTCTTGGCCCTGCTGCTGACCATACCA 60
Db 1 ATGTCGGTTCCTTCTATCTATCTCTTGGCCCTGCTGCTGACCATACCA 60

```

```

Qy 61 GCTTCGCTTATGAAGTGGCAACGTGTCCGGGGTGTACCATGTACGAACGACTGCTCC 120
Db 61 GCTTCGCTTATGAAGTGGCAACGTGTCCGGGGTGTACCATGTACGAACGACTGCTCC 120
Qy 121 AACTCAAGCATAGTGTATGAGGAGGAGCATGATCATGACACCCCGGGTGGTGCC 180
Db 121 AACTCAAGCATAGTGTATGAGGAGGAGCATGATCATGACACCCCGGGTGGTGCC 180
Qy 181 TCGTTTCGGGAGGGCAACTCCTCCCGTTCGTGGGTGGCGCTCACTCCACGCTCGGGCC 240
Db 181 TCGTTTCGGGAGGGCAACTCCTCCCGTTCGTGGGTGGCGCTCACTCCACGCTCGGGCC 240
Qy 241 AGGAACGCCAGGCTCCCAACAGCAATACAGACGCCAGTCGATTGCTTGTGGGGCT 300
Db 241 AGGAACGCCAGGCTCCCAACAGCAATACAGACGCCAGTCGATTGCTTGTGGGGCT 300
Qy 301 GCTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTGGGATCTGTTTCCCTTTTCC 360
Db 301 GCTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTGGGATCTGTTTCCCTTTTCC 360
Qy 361 CAGCTGTTTCACTTCTCAGCTCGCCGGCATCAACAGTAGTACAGACTGCAACTGCTCAATC 420
Db 361 CAGCTGTTTCACTTCTCAGCTCGCCGGCATCAACAGTAGTACAGACTGCAACTGCTCAATC 420
Qy 421 TATCCCGGCATGATCAGGTACCGCATGGCTTGGGATATGATGAACTGGTAATA 479
Db 421 TATCCCGGCATGATCAGGTACCGCATGGCTTGGGATATGATGAACTGGTAATA 479

RESULT 4
AAL48916
ID AAL48916 standard; DNA; 483 BP.
XX AC AAL48916;
XX DT 24-OCT-2002 (first entry)
XX DE Hepatitis C virus clone HCC112A E1 protein coding sequence.
XX KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX virucide; immunostimulant; vaccine; ds.
XX OS Hepatitis C virus.
XX PN WO200255548-A2.
XX PD 18-JUL-2002.
XX PF 11-JAN-2002; 2002WO-EP00219.
XX PR 11-JAN-2001; 2001US-260699P.
XX PR 30-AUG-2001; 2001US-315768P.
XX PA (INNO-) INNOGENETICS NV.
XX PI Maertens G, Bosman F, Buyse M;
XX WPI; 2002-599657/64.
XX P-PSDB; AAO18663.
XX PT New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection
XX Example 2; Page 165-166; 243pp; English.
XX CC The present invention relates to new therapeutic vaccine compositions for
XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
XX composition containing at least one purified recombinant HCV single or
XX specific oligomeric recombinant envelope proteins selected from an E1 and
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are

```


CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.

XX Sequence 483 BP; 85 A; 152 C; 123 G; 123 T; 0 other;

Query Match 98.8%; Score 474.2; DB 24; Length 483;
Best Local Similarity 99.4%; Pred. No. 9.4e-129; Indels 0; Gaps 0;
Matches 476; Conservative 0; Mismatches 3;

QY 1 ATGTCGGTGTCTTCTCTATCTTCTCTTGGCCCTGCTGCTCTGACCATACCA 60
DB 1 ATGCCCCGGTGTCTTCTCTATCTTCTTGGCCCTGCTGCTCTGACCATACCA 60
QY 61 GCTTCGGCTTATGAAGTGCACACGTTGCGGGGTGTACATGTACAGAACTGCTCC 120
DB 61 GCTTCGGCTTATGAAGTGCACACGTTGCGGGGTGTACATGTACAGAACTGCTCC 120
QY 121 AACTCAAGCATGTATGATGAGGAGCGGACATCATGCACACCCCGGTGCGTGC 180
DB 121 AACTCAAGCATGTATGATGAGGAGCGGACATCATGCACACCCCGGTGCGTGC 180
QY 181 TGCCTTCGGGAGGCAACTCTCTCCGTTGTGGGTGGCGTCACTCCACGCTCGCGGC 240
DB 181 TGCCTTCGGGAGGCAACTCTCTCCGTTGTGGGTGGCGTCACTCCACGCTCGCGGC 240
QY 241 AGGAACGCGAGCTCCCAACAGCAATACGAGCGCACCGTGCATTTGCTGTTGGGCT 300
DB 241 AGGAACGCGAGCTCCCAACAGCAATACGAGCGCACCGTGCATTTGCTGTTGGGCT 300
QY 301 GCTGCTTCTCTCCGTTGTAGTGTGGGGATCTCTGCGGATCTGTTTCTTGTTC 360
DB 301 GCTGCTTCTCTCCGTTGTAGTGTGGGGATCTCTGCGGATCTGTTTCTTGTTC 360
QY 361 CAGCTGTTTCACTTCTCACTTCGCGGATCAAAACAGTACGAGCTCACTGCTCAATC 420
DB 361 CAGCTGTTTCACTTCTCACTTCGCGGATCAAAACAGTACGAGCTCACTGCTCAATC 420
QY 421 TATCCCGCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGATGATGATGAT 479
DB 421 TATCCCGCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGATGATGATGAT 479

RESULT 5
AAQ24467
ID AAQ24467 standard; DNA; 1880 BP.

XX AAQ24467;

XX 09-NOV-1992 (first entry)

XX NANB hepatitis virus strain HC-J4 genome.

XX non-A, non-B hepatitis virus; NANBH; PCR; amplification
KW polymerase chain reaction; vaccine; antibody; ss.

XX Non-A, non-B hepatitis virus.

XX Key Location/Qualifiers
FH 342..1880
FT /*tag= a
FT FT /label= HC-J4

XX EP485209-A.

XX 13-MAY-1992.

XX 07-NOV-1991; 91EP-0310297.

XX

PR 08-NOV-1990; 90JP-0304405.

XX (IMMO) IMMUNO JAPAN INC.

XX Nakamura T, Okamoto H;

XX WPI; 1992-160959/20.

XX P-PSDB; AAR24087.

XX Recombinant cDNA of NANBH virus strain HC-J5 and corresp.
PT peptides - useful for diagnosis and in vaccines and immunological
PT pharmaceuticals

XX Disclosure; Page 11; 42pp; English.

XX This sequence is the genome of the non-A, non-B hepatitis virus
CC (NANBH) strain HC-J4. This sequence was derived by amplification
CC by polymerase chain reaction. The nucleotide sequences derived from
CC this amplification can be used to detect NANBH infection which could
CC not be detected by conventional methods. The detection kits allow
CC highly specific and sensitive detection at an early phase of
CC infection. The polypeptide product of this coding sequence can be used
CC for the manufacture of vaccines and immunological pharmaceuticals
CC and also to produce antibodies specific to NANBH.

XX Sequence 1880 BP; 333 A; 593 C; 556 G; 398 T; 0 other;

Query Match 90.2%; Score 432.8; DB 13; Length 1880;
Best Local Similarity 94.3%; Pred. No. 2e-116;
Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 2 TGTCCGGTGTCTTCTCTATCTTCTTGGCCCTGCTGCTCTGACCATACCA 61
DB 847 TGCCTGGTGTCTTCTCTATCTTCTTGGCTTGTCTGCTGCTGACCATACCA 906
QY 62 CTTCCGCTTATGAAGTGCACACGTTGCGGGGTGTACATGTACAGAACTGCTCA 121
DB 907 CTTCCGCTTATGAAGTGCACACGTTGCGGGGTGTACATGTACAGAACTGCTCA 966
QY 122 ACTCAAGCATGTATGATGAGGAGCGGACATCATGCACACCCCGGTGCGTCC 181
DB 967 ACTCAAGCATGTATGATGAGGAGCGGACATCATGCACATCTCCCGGTGCGTCC 1026
QY 182 GGTTCGGAGGCACTCTCCGTTGCTGGGTGCGCTCACTCCACGCTCGCGGCA 241
DB 1027 GCGTTCGGAGGACAAACAGCTCCGTTGCTGGGTAGCGTCACTCCACGCTCGCGGCA 1086
QY 242 GGAACGCCAGCGTCCCAACACAAATACGACGCCACGTCGACTTGTCTGGGCTG 301
DB 1087 GGAATGCCAGCGTCCCACTACGACAAATACGACGCCACGTCGACTTGTCTGGGCGG 1146
QY 302 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTC 361
DB 1147 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTC 1206
QY 362 AGCTTTTCACTTCTCACTTCGCGGATCAAAACAGTACAGGACTGCAACTGCTCAATCT 421
DB 1207 AGCTTTTCACTTCTCGCTTCGCGGATGAGACAGTGCAGGACTGCAACTGCTCAATCT 1266
QY 422 ATCCCGCCATGTATCAGGTACCGCATGCTTGGGATATGATGATGATGATGATGAT 477
DB 1267 ATCCCGCCATTTATCAGGTACCGCATGCTTGGGATATGATGATGATGATGATGAT 1322

RESULT 6
ABA03491
ID ABA03491 standard; DNA; 2187 BP.

XX ABA03491;

XX 15-MAR-2002 (first entry)

XX Cuticle protein 1 and 2 secreting hepatitis C virus related DNA #1.

XX Cuticle protein 1; cuticle protein 2; hepatitis C virus; ds.
XX Unidentified.
XX Key Location/Qualifiers
XX CDS 1..2187
XX /tag= a
XX /product= "AAM47264"
XX /partial
XX /note= "no stop codon"
XX
XX KR97065713-A.
XX
XX 13-OCT-1997.
XX
XX 19-MAR-1996; 96KR-0007404.
XX
XX 19-MAR-1996; 96KR-0007404.
XX
XX (GLDS) LG CHEM LTD.
XX
XX Choo SH, Lee IH, Ryoo WS;
XX PI
XX WPI; 1998-492654/42.
XX P-PSDB; AAM47264.
XX
XX
XX Cuticle protein 1 and 2 secreting hepatitis C virus (Japanese)
XX NoAbstract -
XX
XX Disclosure; Page 2-4; 7pp; Korean.
XX
XX The present invention relates to cuticle protein 1 and 2 secreting
XX hepatitis C virus. The present sequence is a coding sequence
XX provided in the exemplification of the invention.
XX
XX Sequence 2187 BP; 406 A; 669 C; 631 G; 481 T; 0 other;
XX
XX Query Match 90.2%; Score 432.8; DB 19; Length 2187;
XX Best Local Similarity 94.3%; Pred. No. 2.1e-116;
XX Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
XX
QY 2 TGTCCGGTGTCTTTCTCTATCTTCTCTGTGCGCCCTGCTGCTGTGACCATACCAG 61
DB 506 TGCCTGGTGTCTTTCTCTATCTTCTCTGTGCGCCCTGCTGCTGTGACCATACCAG 565
QY 62 CTTCCGCTTATGAAGTCCCAACGCTGTCGGGGTGTACCATGTCCAGAACGACTGTCCA 121
DB 566 CTTCCGCTTATGAAGTCCCAACGCTGTCGGGGTGTACCATGTCCAGAACGACTGTCCA 625
QY 122 ACTCAAGCATAGTGTATGAGGCGGACATGATCATGCACACCCCGGTGCGTCCCT 181
DB 626 ACTCAAGCATAGTGTATGAGGCGGACATGATCATGCACATCTCCGGTGTGCGCT 685
QY 182 GCCTTCGGAGGCGCAACTCTCCGCTGCTGGGTGCGCTCACTCCACGCTCGCGCCA 241
DB 686 GCCTTCGGAGGCGCAACAGCTCCGCTGCTGGGTGCGCTCACTCCACGCTCGCGCCA 745
QY 242 GGAACGCGAGCGTCCCAACACATACGACGCCACGTCGATTTGCTGTTGGGGCTG 301
DB 746 GGAATGCCAGCGTCCCAACATACGACGCCACGTCGATTTGCTGTTGGGGCTG 805
QY 302 CTGCTTTCTGTTCCGCTATGATGAGGCGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
DB 806 CTGCTTTCTGTTCCGCTATGATGAGGCGGATCTCTGCGGATCTGTTTCTGTTTCCC 865
QY 362 AGCTGTTCACCTTCTCACTCGCGGCGATCAAAACAGTACAGGACTGCAACTGCTCAATCT 421
DB 866 AGCTGTTCACCTTCTCACTCGCGGCGATGAGACAGTACAGGACTGCAACTGCTCAATCT 925
QY 422 ATCCCGGCCATGATACAGGTACCGCATGCGCTTGGGATATGATGATGAAGTGTAA 477
DB 926 ATCCCGGCCATGATACAGGTACCGCATGCGCTTGGGATATGATGATGAAGTGTAA 981

RESULT 7
AAQ43889
ID AQA43889 standard; cDNA to mRNA; 2540 BP.
XX
XX AQA43889;
XX
XX 21-OCT-1993 (first entry)
XX
XX NANB hepatitis virus polynucleotide N-2540-2.
XX
XX Non-A, non-B; virus; polymerase chain reaction; detection;
XX sensitive; specific; HCV; NANBH; ss.
XX
XX Non-A, non-B hepatitis virus.
XX
XX Key Location/Qualifiers
XX CDS 342..2540
XX /tag= a
XX 5'UTR 1..341
XX /tag= b
XX /note= "from 5' terminal of NANBH virus RNA"
XX
XX JP05091884-A.
XX
XX 16-APR-1993.
XX
XX 10-APR-1991; 91JP-0196175.
XX
XX 12-JUN-1990; 90JP-0153401.
XX
XX 08-NOV-1990; 90JP-0304405.
XX
XX (NAKA/) NAKAMURA T.
XX
XX WPI; 1993-199637/25.
XX P-PSDB; AAR38279.
XX
XX Antigen related to non-A and non-B hepatitis virus - comprises
XX non-translation region comprising 340 - 341 mols. of nucleotides,
XX non-translation region comprising 1885 - 2551 mols. of
XX nucleotides including region 1,149 and, etc.
XX
XX Claim 3; Page 19-20; 73pp; Japanese.
XX
XX The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
XX which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
XX The polypeptide it encodes may be used in a system for detecting
XX NANB hepatitis. This method is highly specific and sensitive, and
XX can detect NANB hepatitis virus which could not be detected by
XX conventional methods.
XX
XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;
XX
XX Query Match 90.2%; Score 432.8; DB 14; Length 2540;
XX Best Local Similarity 94.3%; Pred. No. 2.2e-116;
XX Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
XX
QY 2 TGTCCGGTGTCTTTCTCTATCTTCTCTGTGCGCCCTGCTGCTGTGACCATACCAG 61
DB 847 TGCCTGGTGTCTTTCTCTATCTTCTCTGTGCGCCCTGCTGCTGTGACCATACCAG 906
QY 62 CTTCCGCTTATGAAGTCCCAACGCTGTCGGGGTGTACCATGTCCAGAACGACTGTCCA 121
DB 907 CTTCCGCTTATGAAGTCCCAACGCTGTCGGGGTGTACCATGTCCAGAACGACTGTCCA 966
QY 122 ACTCAAGCATAGTGTATGAGGCGGACATGATCATGCACACCCCGGTGCGTCCCT 181
DB 967 ACTCAAGCATAGTGTATGAGGCGGACATGATCATGCATCTCCCGGTGCGTCCCT 1026
QY 182 GCCTTCGGAGGCGCAACTCTCCGCTGCTGGGTGCGCTCACTCCACGCTCGCGCCA 241
DB 1027 GCCTTCGGAGGCGCAACAGCTCCCGTGTGTTGGGTAGCGCTCACTCCACGCTCGCGCCA 1086

FH Key Location/Qualifiers
 FT 342..9374
 FT CDS /*tag= a
 PN WO904008-A2.
 XX 28-JAN-1999.
 XX 16-JUL-1998; 98WO-US14688.
 XX 27-JAN-1998; 98US-0014416.
 PR 18-JUL-1997; 97US-0053062.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Bukh J, Emerson SU, Purcell RH, Yanagi M;
 PI WPI; 1999-132252/11.
 DR P-PSDB; AAW98022.
 XX New isolated hepatitis C virus nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of HCV
 PT infections and for developing screening assays
 XX Claim 3; Fig 14A-F; 126pp; English.
 XX The present sequence comprises the nucleic acid sequence of the
 CC genome of infectious hepatitis C virus (HCV) genotype 1b strain
 CC HC-J4 (ATCC 209596) that is capable of expressing this virus when
 CC transfected into cells. HC-J4 was obtained from acute phase plasma
 CC of a chimpanzee experimentally infected with serum containing
 CC HC-J4/91. The claimed infectious nucleic acid sequence can be used
 CC to produce chimeric genomes (see AAX24833) consisting of the open
 CC reading frames of infectious nucleic acid sequences of other
 CC genotypes (including genotypes 1-6) and subtypes (such as 1b, 2a,
 CC 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also relates to
 CC the introduction of mutations or deletions into infectious nucleic
 CC acid sequences in order to produce an attenuated HCV virus suitable
 CC for vaccine development. Infectious nucleic acid sequences can
 CC also be used to produce attenuated virus via passage in vitro or in
 CC vivo of the viruses produced by transfection of a host cell with
 CC the infectious nucleic acid sequence. Vaccines comprising one or
 CC more polypeptides made from the infectious nucleic acid sequence are
 CC used to immunise mammals, especially humans, against hepatitis C.
 CC The nucleic acid sequences can also be used to induce protective
 CC immunity against the virus. The nucleic acid sequences or their
 CC encoded proteases (e.g. NS3 protease) can additionally be used to
 CC develop screening assays to identify antiviral agents for HCV.
 XX Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;
 SQ
 Query Match 89.2%; Score 428; DB 20; Length 9595;
 Best Local Similarity 93.7%; Pred. No. 9.1e-115;
 Matches 446; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
 QY 2 TGTCCGGTTCCTTTCTCTATCTCTTGGCCCTCTGTCTGTCTGACCATACAG 61
 DB 847 TGCCCGGTTCCTTTCTCTATCTCTTGGCTCTGTCTGTCTGTTGACCATCCAG 906
 QY 62 CTTCCGCTTATGAAGTGCACACGTGTCCGGGGTGTACATGTACGAACGACTGCTCCA 121
 DB 907 CTTCCGCTTATGAAGTGCACACGTGTCCGGGGTGTACATGTACGAACGACTGCTCCA 966
 QY 122 ACTCAAGCATAGTGTATGAGGAGGAGCATATCATGCACACCCCGGTCGGGCCCT 181
 DB 967 ACTCAAGCATAGTGTATGAGGAGGAGGAGCATATCATGCATCTCCCGGTCGGGCCCT 1026
 QY 182 GCGTTCGGAGGAGCAACTCTCTCCGTTGCTGGGGTGGCGTCACTCCACGCTCGCGGCA 241
 DB 1027 GTGTTAGAGGGGTACACGCTCCCGTTGCTGGGTAGCGTCACTCCACGCTCGCGGCA 1086
 QY 242 GGAACGCCAGCTCCCAACACGACATAGAGCGCCAGTCGATTTGCTGTTGGGGTGTG 301

Db 1087 GGAATGCCAGGTCCTCCCACTAGCAACAATAGCAGCCACCGCTGCTGTTGGAGCGG 1146
 QY 302 CTGCTTTCTGTTCCGCTATGCTAGCTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 361
 Db 1147 CTGCTTTCTGCTCCGCTATGCTAGCTGGGGATCTCTCGGATCTATTTCTGCTCTCC 1206
 QY 362 AGCTGTTTCACTTTCTCACCTCGCCGCGCATCAACAGTACAGGACTCAACTGCTCAATCT 421
 Db 1207 AGCTGTTTCACTTTCTCGCCTCGCGCATGAGACAGTGCAGGACTGCAACTGCTCAATCT 1266
 QY 422 ATCCCGCCCATGATGAGGTACCGCATGCGTGGGATGATGATGATGATGATGATGATGATGAT 477
 Db 1267 ATCCCGCCCATGATGAGGTACCGCATGCGTGGGATGATGATGATGATGATGATGATGATGAT 1322
 RESULT 12
 AAC86939
 ID AAC86939 standard; DNA; 9595 BP.
 XX
 AC AAC86939;
 XX
 DT 02-APR-2001 (first entry)
 XX Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 1b.
 DE
 XX Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
 KW HCV; vaccine; viral inhibitor; antiviral; ss.
 XX Hepatitis C virus.
 OS
 XX Key Location/Qualifiers
 FT CDS 342..9374
 FT /*tag= a
 XX WO200075352-A2.
 XX 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-US15527.
 XX 04-JUN-1999; 99US-0137817.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Nam J, Bukh J, Emerson SU, Purcell RH;
 PI WPI; 2001-071081/08.
 DR P-PSDB; AAB31170.
 XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus
 PT genome in which the (non-)structural region has been replaced by
 PT hepatitis C virus (HCV) genome useful for treating or preventing HCV
 PT signs and symptoms -
 XX Disclosure; Fig 4A-F; 97pp; English.
 XX The specification describes a nucleic acid comprising a chimeric virus
 CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
 CC the (non-)structural region has been replaced by the (non-)structural
 CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
 CC the chimeric virus and the chimeric virus are useful for identifying
 CC cell lines capable of supporting the replication of these chimeric
 CC viruses, in screening for neutralizing antibodies to HCV of different
 CC genotypes, in the production of HCV-BVDV virions for the development
 CC of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal,
 CC in studying the molecular properties of HCV indirectly in vitro, and in
 CC identifying inhibitors of viral enzyme activity which would be useful
 CC as antiviral agents. Formulations or compositions comprising the
 CC chimeric virions may be used to treat or prevent the signs and symptoms
 CC of HCV. The present sequence represents a HCV clone, which is used
 CC to construct chimeric nucleic acids of the invention.
 XX Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;

Qy 422 ATCCGGCCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGAACCTGGTAA 477

Db 482 ATCTTGGCCACGTATCAGGTATCGCATGGCTTGGGATATGATGATGAACCTGGTCA 537

Search completed: December 19, 2003, 18:51:06
Job time : 133.7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1234.83 Seconds

(without alignments)

9447.586 Million cell updates/sec

Title: US-09-899-303A-11

Perfect score: 480

Sequence: 1 ATGTCGGTGTCTCTTCTC.....TGATGATGAACCTGGTAATAG 480

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_nam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	65.8	13.7	488	9	AV755731
C 2	54.6	11.4	492	9	AV758366
C 3	43.2	9.0	534	14	CD040840
C 4	41.6	8.7	664	29	BZ645446

C	5	40.6	8.5	526	9	AL825643
	6	40.2	8.4	399	9	AV638521
	7	40.2	8.4	434	9	AV637507
	8	40.2	8.4	440	9	AV637983
	9	40.2	8.4	450	9	AV637259
	10	40.2	8.4	451	9	AV637328
	11	40.2	8.4	451	9	AV637643
	12	40.2	8.4	453	9	AV634724
	13	40.2	8.4	454	9	AV637050
	14	40.2	8.4	456	9	AV635382
	15	40.2	8.4	473	9	AV632765
	16	40.2	8.4	481	9	AV635503
	17	40.2	8.4	485	9	AV632811
	18	40.2	8.4	506	9	AV392445
	19	40.2	8.4	508	9	AV634095
	20	40.2	8.4	526	9	AV641895
	21	40.2	8.4	533	9	AV638125
	22	40.2	8.4	537	9	AV632335
	23	40.2	8.4	588	9	AV387329
	24	40	8.3	1186	13	BX421743
	25	39.4	8.2	1201	9	AL565958
C	26	39.2	8.2	624	14	CD206870
C	27	39.2	8.2	656	14	CB924688
	28	39	8.1	497	9	AV633658
	29	39	8.1	610	14	CB657655
	30	39	8.1	856	29	BZ578381
	31	39	8.1	872	29	BZ555011
C	32	38.4	8.0	645	29	CNS01213
C	33	38	7.9	771	29	BZ530934
C	34	37.8	7.9	490	9	AV634529
C	35	37.8	7.9	705	14	CA618797
C	36	37.8	7.9	1039	13	BX415186
C	37	37.6	7.8	309	12	BI098866
C	38	37.6	7.8	431	9	AV639153
C	39	37.4	7.8	394	12	BZ209795
C	40	37.4	7.8	501	9	AV638474
C	41	37.2	7.8	431	9	AV641448
C	42	37.2	7.8	792	13	BX391120
C	43	37	7.7	910	29	CNS0060N
	44	36.8	7.7	214	12	BM686105
	45	36.8	7.7	913	14	CA487901

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

488 bp mRNA linear EST 19-OCT-2000
AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
AV755731.1 GI:10913579
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 488)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES

us-09-899-303a-11.rst

Mon Dec 22 13:28:34 2003

```

source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN
Query Match 13.7%; Score 65.8; DB 9; Length 488;
Best Local Similarity 67.0%; Pred. No. 2.3e-07;
Matches 124; Conservative 0; Mismatches 57; Indels 4; Gaps 2;

QY 292 GTTGGGCTGCTGCTTCTGTTCCGTATGTACGTGGGGATCTCTGGGATCTGTTTC 351
Db 472 GTGGTGTACACTCGCTCTCTAGCTGTGGGACCTCTGGACGAGTGATG 413
QY 352 CTTGTTTCCAGCTGTTTCACTTCTCAGCTCGCGGCATCAACAGTACAGACTGCAAC 411
Db 412 CTTGCAGTTTCAGCTGATCA---TCTGGCCTCAGCACCATGAGTTTGTGATGATGCAAC 356
QY 412 TGCTCAATCTATCCGGCCATGTATCAGTTCACCGCATG-GCTTGGGATATGATGAA 470
Db 355 TGCTCCATCTATCTGGGGCCATCACTGGACACCGTATGAGCATGGGACATGATGAA 296
QY 471 CTGGT 475
Db 295 CTGGT 291

RESULT 2
AV758366/c 492 bp mRNA linear EST 19-OCT-2000
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. .492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN

```

```

Query Match 11.4%; Score 54.6; DB 9; Length 492;
Best Local Similarity 65.7%; Pred. No. 0.00027;
Matches 111; Conservative 0; Mismatches 54; Indels 4; Gaps 2;

QY 308 TCTGTTCCGCTATGTACGTGGGGATCTCTGGGATCTCTGTTTCTTCCAGCTGT 367
Db 457 TGTGATCAGTCTACTAGCTGTGGACCTCTGCGTTGGGTATCGCTTGCAGCCCACTGA 398
QY 368 TCACCTTCTCCTCGCGGCATCAACAGTACAGACTGCAACTGCTCAATCTATCCCG 427
Db 397 TTA---TCTCTCAGCAGCAACATTTGGTTTGTGCAAGATGCAACTGCTCATCTATCTG 341
QY 428 GCCATGTATCAGTTC-ACCGCATGGCTTGGGATATGATGATGAACTGGT 475
Db 340 GCTGCATCACTGGACACTACAGTATGGCATAGCTATGATGATGAACTGGT 292

RESULT 3
CD040840 534 bp mRNA linear EST 09-MAY-2003
LOCUS pHB036xB09f_300663 psHB: Infected hypocotyl soybean host. 48 hrs
DEFINITION post infection Phycophthora sojae cDNA clone sHB036B09 5, mRNA
sequence.
ACCESSION CD040840
VERSION CD040840.1 GI:30502701
KEYWORDS EST.
Phycophthora sojae
Phycophthora sojae
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
ORGANISM Phycophthora.
1 (bases 1 to 534)
Tyler, B.M., Judelson, H.S., Gijzen, M., Dean, R.A. and Waugh, M.E.
USDA-IPAFS: Expression of Phycophthora sojae genes during infection
and propagation
Unpublished
Contact: Tyler B
Tyler lab
VBI
1880 Pratt Dr., Blacksburg, VA 24061, USA
Tel: 540-231-7318
Email: bmtyley@vt.edu
PCR Primers
FORWARD: BK reverse
Plate: 036 row: B column: 09
Seq primer: BK reverse
High quality sequence stop: 534.
Location/Qualifiers
1. .534
/organism="Phycophthora sojae"
/mol_type="mRNA"
/db_xref="taxon:67593"
/clone="sHB036B09"
/tissue_type="infected host tissue"
/cell_line="p6497"
/dev_stage="48 hour post infection"
/clone_lib="psHB: Infected hypocotyl soybean host. 48 hrs
post infection"
/note="Vector: psB-CMV; Site_1: EcoRI; Site_2: XhoI;
USDA-IPAFS: Expression of Phycophthora sojae genes during
infection and propagation."
BASE COUNT 101 a 187 c 159 g 87 t
ORIGIN
Query Match 9.0%; Score 43.2; DB 14; Length 534;
Best Local Similarity 47.4%; Pred. No. 0.36;
Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 70 TATGAAGTGGCAACGTGTCGGGGTGTACCATGTACAGCAAGACTGCTCAACTAAGC 129
Db 200 TACGGCTGGCGGAGATTACGTATCCGATGGCTTCGCCGCTTCTACAACTGGACC 259
QY 130 ATAGTGTATGAGGAGCGGACATGATCATGCACACCCCCGGGTGCGTTCGGTTCGG 189

```



```

Db      260 TCATGGACAGAGAGGCGCCGATCATGCTGACCCCAAGACGGTGGCCAACTTCAC 319
Qy      190 GAGGGCAACTCTCCGCTTCTGGGTGGGGCTCACTCCCAACGCTCGCGGCCAGGAACGCG 249
Db      320 CACTACGGCGGCACCATCTCGCTCGACCGTGGTGGCTTCGACGTGGACAAGATTATC 379
Qy      250 AGGTTCCCAACAGCAATACAGCCAGCTGATTTGCTGTTGGGGCTGCTGCTTTTC 309
Db      380 AACTTCTTACGACGAAACCGGCTCTCGAGGTATACGTATCGTCGCGGTGACGGCACCCAC 439
Qy      310 TGTTCGCTATGTACGTGGGGATCTCTGCGG 341
Db      440 CGTCCGCCCAACAGATCTCGAGGAGTGC 471

RESULT 4
LOCUS   BZ645446/c
DEFINITION BZ645446.1 Zea mays genomic clone ZMBMa0133P04, genomic survey sequence.
ACCESSION BZ645446
VERSION   BZ645446.1 GI:28107610
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays

REFERENCE
AUTHORS   Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Numborg, A., Robbins, D., and Lakey, N.
TITLE     Consortium for Maize Genomics
JOURNAL   Unpublished
COMMENT    Other GSSs: OGCBJ86TM
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source
1..664
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0133P04"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

BASE COUNT 103 a 238 c 222 g 101 t
ORIGIN

Query Match 8.7%; Score 41.6; DB 29; Length 664;
Best Local Similarity 50.5%; Pred. No. 1.1;
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy      119 CCAACTCAAGCAPAGTGTATGAGCGACGCGACATGATCATGCACACCCCGGTGGGTGC 178
Db      266 CCGACGAGACGACACGTGGTGGCGCCCAAGACGAGGTGCGGACGAGGCGCGGAC 207
Qy      179 CTGCGCTTGGGAGGGAACCTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGG 238
Db      206 CGTACCTGGGTGGCGAGCTCCACCAAGTGTGAGAAGCCCTGCGCTGCTGTTGCGG 147
Qy      239 CCAGGAACCGCGTCCCAACATACGACGCCACGTCGATTCTGCTGTTGGG 298
Db      146 ACAGGAAGGCCCGCTTCCACGACGACCCACCAGCGGCTACCCCTTCTCGGTGGT 87

```

```

Qy      299 CTGCTGCTTTCTGTTCCGCT 318
Db      86 GTGGTGGTGGTGGTGGTGGT 67

RESULT 5
LOCUS   AL825643/c
DEFINITION AL825643 p:234 Triticum aestivum cDNA clone A09_p234_plate_14, mRNA sequence.
ACCESSION AL825643
VERSION   AL825643.1 GI:21837164
KEYWORDS  EST.
SOURCE    Triticum aestivum (bread wheat)
ORGANISM  Triticum aestivum

REFERENCE
AUTHORS   Wilson, I., Beswick, R., Shepherd, S., Barker, G., Parker, J., Owen, P., Edwards, D., Coghill, J., Holdsworth, M., Lenton, J., Shewry, P., and Edwards, K.
TITLE     A BSRRC-funded wheat EST resource for the academic community
JOURNAL   Unpublished
COMMENT    Contact: Barker G
Institute of Arable Crop Research
Long Ashton, Bristol BS41 9AF United Kingdom.

FEATURES
source
1..526
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="mercia"
/db_xref="taxon:4565"
/clone="A09_p234_plate_14"
/tissue_type="drought stressed seedlings"
/dev_stage="21 days old"
/clone_lib="p:234"

BASE COUNT 97 a 179 c 164 g 86 t
ORIGIN

Query Match 8.5%; Score 40.6; DB 9; Length 526;
Best Local Similarity 57.5%; Pred. No. 1.8;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy      150 CATGATCATGCACACCCCGGTGCGTGCCTCGCTCGGAGGCAACTCTCCCGTTG 209
Db      151 CTCTCTCCCGAACCCCGGGTGGATCAGTCGCGAGTTCTTCGCGAGCCCTTCTTGGC 92
Qy      210 CTGGGTGGCGCTCACTCCACGCTCGCGCCAGGAACGCCAGCGTCCCAACGACAAT 269
Db      91 CTGGGCGGCGCGCGGCGGCGGTCGCGACAGGAACACAGCACCGCGCGCGCGGAC 32
Qy      270 ACACGC 276
Db      31 CACACCC 25

RESULT 6
LOCUS   AV638521
DEFINITION AV638521 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii cDNA clone HC087d07_r 5', mRNA sequence.
ACCESSION AV638521
VERSION   AV638521.1 GI:10781841
KEYWORDS  EST.
SOURCE    Chlamydomonas reinhardtii
ORGANISM  Chlamydomonas reinhardtii

REFERENCE
AUTHORS   Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.
TITLE     1 (bases 1 to 399)
Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y., and Tabata, S.
Generation of expressed sequence tags from low-CO2 and high-CO2

```


JOURNAL	adapted cells of Chlamydomonas reinhardtii
MEDLINE	DNA Res. 7 (5), 305-307 (2000)
PUBMED	20539644
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/. Location/Qualifiers
FEATURES	1..399 /organism="Chlamydomonas reinhardtii" /mol_type="mRNA" /db_xref="taxon:3055" /clone_lib="HC074a01_r" /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"
BASE COUNT	70 a 144 c 122 g 63 t
ORIGIN	
Query Match	8.4%; Score 40.2; DB 9; Length 399;
Best Local Similarity	51.4%; Pred. No. 2.2;
Matches	93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY	30 CTTGGCCCTGCTGTCTGTCATACCAAGTTCGCCCTTATGAAGTGCGCAACGTGC 89
DB	
Db	44 CTTGACCCCCCTGGACGGCTCGTCATCGTCGACCAACTTCGCGTGGGCACCATCTT 103
QY	90 CGGGGTGTAACAAGTTCACGAACGACTGCTCCAACCTCAAGCATAAGTAGTATGAGGCGGGA 149
DB	
Db	104 CGCGGTGTGCCCGGCGACAAAGTGCACCAACTCACCGCGCGGACGAGGTGGTGC 163
QY	150 CATGATCATGCACACCCCGGTCGTCGCCCTCGCTCGGGAGGCGCAACTCTCCCGTTG 209
DB	
Db	164 CATGGCATCTACGGTCCCGGACCCTGTTCTGCATTGCCCTGAAGGACGCCCGCGCTG 223
QY	210 C 210
Db	224 C 224
RESULT 7	
AV637507	434 bp mRNA linear EST 15-DEC-2000
LOCUS	AV637507 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION	CDNA clone HC074a01_r 5', mRNA sequence.
ACCESSION	AV637507.1 GI:10780827
VERSION	EST.
KEYWORDS	Chlamydomonas reinhardtii
SOURCE	Chlamydomonas reinhardtii
ORGANISM	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.
REFERENCE	Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y. and Tabata,S. Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
AUTHORS	DNA Res. 7 (5), 305-307 (2000)
TITLE	Adapted cells of Chlamydomonas reinhardtii
JOURNAL	DNA Res. 7 (5), 305-307 (2000)
MEDLINE	20539644
PUBMED	11089912
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/. Location/Qualifiers
FEATURES	1..434 /organism="Chlamydomonas reinhardtii"
source	
BASE COUNT	76 a 156 c 137 g 71 t
ORIGIN	
Query Match	8.4%; Score 40.2; DB 9; Length 434;
Best Local Similarity	51.4%; Pred. No. 2.2;
Matches	93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY	30 CTTGGCCCTGCTGTCTGTCATACCAAGTTCGCCCTTATGAAGTGCGCAACGTGC 89
DB	
Db	21 CTTGACCCCCCTGGACGGCTCGTCATCGTCGACCAACTTCGCGTGGGCACCATCTT 80
QY	90 CGGGGTGTAACAAGTTCACGAACGACTGCTCCAACCTCAAGCATAAGTAGTATGAGGCGGGA 149
DB	
Db	81 CGCGGTGTGCCCGGCGACAAAGTGCACCAACTCACCGCGCGGACGAGTGGTGC 140
QY	150 CATGATCATGCACACCCCGGTCGTCGCCCTCGCTCGGGAGGCGCAACTCTCCCGTTG 209
DB	
Db	141 CATGGCATCTACGGTCCCGGACCCTGTTCTGCATTGCCCTGAAGGACGCCCGCGCTG 200
QY	210 C 210
Db	201 C 201
RESULT 8	
AV637983	440 bp mRNA linear EST 15-DEC-2000
LOCUS	AV637983 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION	CDNA clone HC080C04_r 5', mRNA sequence.
ACCESSION	AV637983
VERSION	AV637983.1 GI:10781303
KEYWORDS	EST.
SOURCE	Chlamydomonas reinhardtii
ORGANISM	Chlamydomonas reinhardtii
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.
AUTHORS	Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y. and Tabata,S. Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
TITLE	Adapted cells of Chlamydomonas reinhardtii
JOURNAL	DNA Res. 7 (5), 305-307 (2000)
MEDLINE	20539644
PUBMED	11089912
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/. Location/Qualifiers
FEATURES	1..440 /organism="Chlamydomonas reinhardtii"
source	
BASE COUNT	76 a 156 c 137 g 71 t
ORIGIN	
Query Match	8.4%; Score 40.2; DB 9; Length 434;
Best Local Similarity	51.4%; Pred. No. 2.2;
Matches	93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY	30 CTTGGCCCTGCTGTCTGTCATACCAAGTTCGCCCTTATGAAGTGCGCAACGTGC 89
DB	
Db	21 CTTGACCCCCCTGGACGGCTCGTCATCGTCGACCAACTTCGCGTGGGCACCATCTT 80
QY	90 CGGGGTGTAACAAGTTCACGAACGACTGCTCCAACCTCAAGCATAAGTAGTATGAGGCGGGA 149
DB	
Db	81 CGCGGTGTGCCCGGCGACAAAGTGCACCAACTCACCGCGCGGACGAGTGGTGC 140
QY	150 CATGATCATGCACACCCCGGTCGTCGCCCTCGCTCGGGAGGCGCAACTCTCCCGTTG 209
DB	
Db	141 CATGGCATCTACGGTCCCGGACCCTGTTCTGCATTGCCCTGAAGGACGCCCGCGCTG 200
QY	210 C 210
Db	201 C 201
mol_type="mRNA"	
/strain="C9"	
/db_xref="taxon:3055"	
/clone_lib="HC074a01_r"	
/clone_lib="Chlamydomonas reinhardtii 5% CO2"	
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"	
BASE COUNT	80 a 149 c 130 g 75 t
ORIGIN	
Query Match	8.4%; Score 40.2; DB 9; Length 434;
Best Local Similarity	51.4%; Pred. No. 2.2;
Matches	93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY	30 CTTGGCCCTGCTGTCTGTCATACCAAGTTCGCCCTTATGAAGTGCGCAACGTGC 89
DB	
Db	21 CTTGACCCCCCTGGACGGCTCGTCATCGTCGACCAACTTCGCGTGGGCACCATCTT 80
QY	90 CGGGGTGTAACAAGTTCACGAACGACTGCTCCAACCTCAAGCATAAGTAGTATGAGGCGGGA 149
DB	
Db	81 CGCGGTGTGCCCGGCGACAAAGTGCACCAACTCACCGCGCGGACGAGTGGTGC 140
QY	150 CATGATCATGCACACCCCGGTCGTCGCCCTCGCTCGGGAGGCGCAACTCTCCCGTTG 209
DB	
Db	141 CATGGCATCTACGGTCCCGGACCCTGTTCTGCATTGCCCTGAAGGACGCCCGCGCTG 200
QY	210 C 210
Db	201 C 201
mol_type="mRNA"	
/strain="C9"	
/db_xref="taxon:	


```

Query Match      8.4%; Score 40.2; DB 9; Length 440;
Best Local Similarity 51.4%; Pred. No. 2.2; Mismatches 0; Indels 0; Gaps 0;
Matches 93; Conservative 0;

QY 30 CTTGGCCCTGCTGCTCTGCTGACCATACAGCTTCCGCTTATGAAGTGCAGCAACGTGTC 89
DB 80 CTTTCGACCCCTGGAGCGCTGCTCCATCGTGACACCAACTTCGCCGTGGCACCATTCTT 139

QY 90 CGGGGTGTACCATGTTCACGAAGCATGTCTCCAACTCAAGCATAGTGTATGAGGACGCGGA 149
DB 140 CGCGCTGTGGCCGCGGACAAAGTGACCAACATCACCGCGCGGACGAGCATGGCTGCCGG 199

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCGTTCGGGAGGGCAACTCTCTCCGTTG 209
DB 200 CATGGCATCTACGGTCCCGCACCCTGTTCTGTCATTGCCCTGAAGGACGCCCCGGCTG 259

QY 210 C 210
DB 260 C 260

RESULT 9
AV637259
LOCUS
DEFINITION
AV637259 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
CDNA clone HC070906_r 5', mRNA sequence.
ACCESSION
AV637259
VERSION
AV637259.1 GI:10780579
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
ORGANISM
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 450)
AUTHORS
Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
PUBMED
11089912
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
1..450
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC070906_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      74 a 153 c 148 g 75 t
ORIGIN
Query Match      8.4%; Score 40.2; DB 9; Length 450;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTGCTCTGCTGACCATACAGTTCGCTTATGAAGTGCAGCAACGTGTC 89
DB 221 CTTTCGACCCCTGGAGCGCTGCTCCATCGTGACACCAACTTCGCCGTGGCACCATTCTT 280

QY 90 CGGGGTGTACCATGTTCACGAAGCATGTCTCCAACTCAAGCATAGTGTATGAGGACGCGGA 149
DB 281 CGCGCTGTGGCCGCGGACAAAGTGACCAACATCACCGCGCGGACGAGCATGGCTGCCGG 340

```

```

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCGTTCGGGAGGGCAACTCTCTCCGTTG 209
DB 341 CATGGCATCTACGGTCCCGCACCCTGTTCTGTCATTGCCCTGAAGGACGCCCCGGCTG 400

QY 210 C 210
DB 401 C 401

RESULT 10
AV637328
LOCUS
DEFINITION
AV637328 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
CDNA clone HC071f12_r 5', mRNA sequence.
ACCESSION
AV637328
VERSION
AV637328.1 GI:10780648
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
ORGANISM
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 451)
AUTHORS
Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
PUBMED
11089912
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
1..451
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC071f12_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      74 a 150 c 150 g 77 t
ORIGIN
Query Match      8.4%; Score 40.2; DB 9; Length 451;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTGCTCTGCTGACCATACAGTTCGCTTATGAAGTGCAGCAACGTGTC 89
DB 178 CTTTCGACCCCTGGAGCGCTGCTCCATCGTGACACCAACTTCGCCGTGGGACCATTCTT 237

QY 90 CGGGGTGTACCATGTTCACGAAGCATGTCTCCAACTCAAGCATAGTGTATGAGGACGCGGA 149
DB 238 CGCGCTGTGGCCGCGGACAAAGTGACCAACATCACCGCGCGGACGAGTGGCTGCCGG 297

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCGTTCGGGAGGGCAACTCTCTCCGTTG 209
DB 298 CATGGCATCTACGGTCCCGCACCCTGTTCTGTCATTGCCCTGAAGGACGCCCCGGCTG 357

QY 210 C 210
DB 358 C 358

RESULT 11
AV637643
LOCUS

```


Generation of expressed sequence tags from low-CO₂ and high-CO₂ adapted cells of *Chlamydomonas reinhardtii*
DNA Res. 7 (5), 305-307 (2000)

Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/>

```

FEATURES
    source
        Location/Qualifiers
            1..453
                /organism="Chlamydomonas reinhardtii"
                /mol_type="mRNA"
                /strain="C9"
                /db_xref="taxon:3055"
                /clone="HC037a06_r"
                /clone_lib="Chlamydomonas reinhardtii 5% CO2"
                /notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
                xhoI; The cDNA library was constructed from cells cultured
                in medium with bubbling air containing 5% carbon

```

BASE COUNT	76 a	154 c	148 g	75 t
ORIGIN				

[illegible]

AV637050	454 bp	mRNA	linear	EST	15-DEC-2000
AV637050	<i>Chlamydomonas reinhardtii</i>	5' CO2	<i>Chlamydomonas reinhardtii</i>		
	3'-GGGTCCTCG-04	5'	mRNA sequence.		

AV637050
AV637050.1 GI:10780370

ESI.
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae. Chlamydomonas.

1 (bases 1 to 454)
Asamizu, E., Miura, K., Kuchō, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata S.
Generation of expressed sequence tags from low-CO₂ and high-CO₂
adapted cells of *Chlamydomonas reinhardtii*
Plant Physiol. 75, 105-107 (1990)

DNA KEY: / (37, 502-504) (2005)
20539644
11089912

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 282-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant>
Location/Qualifiers
1. . 454

FEATURES

AV637643 Chlamydomonas reinhardtii 5' CO2 Chlamydomonas reinhardtii cDNA clone HC075a03 r 5', mRNA sequence.

AV637643
 AV637643.1 GI:10780963
 EST.
 Chlamydomonas reinhardtii
 Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 451)
Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
Nakamura, Y. and Tabata, S.
Generation of expressed sequence tags from low-CO₂ and high-CO₂
adapted cells of *Chlamydomonas reinhardtii*
DNA Res. 7 (5), 305-307 (2000)
20539644
11089912
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research

Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>
Location/Qualifiers

```

1. .451
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone_hc="HC075g03_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
76 a 151 c 149 g 75 t
8.4%; Score 40.2; DB 9; Length 451;

```

90 CCGGGTGTACATGTCACGAACGACTGTCTCAAATCAAGCATAGTGTATGAGGACGCGA 14

256 CCGCGTGTGGCCGCGCAAGCTGACCAATCAACGATCAACCGCCGCGACAGTGGTGGCCGG 315

316 CATGGGATCTACGGTCCCCGCACCGTGTCTGCAATTGCCCTGAAGGACGCCCGCGCTG 375

210 C 210
—
376 C 376

Accession	Length	Source	EST
AV634724	453 bp	linear	EST 15-DEC-2001
AV634724	Chlamydomonas reinhardtii 5% CO2	Chlamydomonas reinhardtii	
	NC022062.5	5' mRNA sequence.	

CDNA clone HC03/a06_1.5; mRNA sequence:

AV634724	
AV634724	1
AV634724	CT·10778044

AV0834/29.1 GR1267/0000
EST.
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae;
Chlamydomonadales; Chlamydomonas.
1 (bases 1 to 453)
Kobayashi, M., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohshima, K.

REFERENCE


```
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/db_strain="C9"
/db_xref="taxon:3055"
/clone="HC068a04_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      82 a 153 c 144 g 75 t
ORIGIN

Query Match      8.4%; Score 40.2; DB 9; Length 454;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTCTGCTGACCATACAGCTTCCGCTTATGAAGTGGCAACGTGTC 89
DB 52 CTTGACCCCTGGAGCGGTCTGTCATCGTCGACACCACTTCGCCGTGGGCAACATCTT 111

QY 90 CGGGGTGTACCATGTCCAGAACGACTGTCTCCAACTCAAGCATAGTGTATGAGGACGCGGA 149
DB 112 CGCGTGTGGCGCGGACAAAGCTGACCAATCACCAGCGCGGAGCAGGTGGCTGCCGG 171

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCTCGGTTCGGAGGGCAACTCTCTCCGTTG 209
DB 172 CATGGGCATCTAGCGTCCCGCAGCCGCTGTTCTGCAATTGCGCTGAAGGACGCCCGCGCTG 231

QY 210 C 210
DB 232 C 232

RESULT 14
AV635382
LOCUS
DEFINITION
AV635382 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
cDNA clone HC045f10_r 5', mRNA sequence.
ACCESSION
AV635382
VERSION
AV635382.1 GI:10778702
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 456)
AUTHORS
Asanizu,E., Miura,K., Kuchro,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL
DNA Res. 7 (5), 305-307 (2000)
MEDLINE
20539644
PubMed
11089912
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
1..456
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/db_strain="C9"
/db_xref="taxon:3055"
/clone="HC045f10_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      75 a 158 c 148 g 75 t
ORIGIN

/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/db_strain="C9"
/db_xref="taxon:3055"
/clone="HC068a04_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      82 a 153 c 144 g 75 t
ORIGIN

Query Match      8.4%; Score 40.2; DB 9; Length 454;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTCTGCTGACCATACAGCTTCCGCTTATGAAGTGGCAACGTGTC 89
DB 52 CTTGACCCCTGGAGCGGTCTGTCATCGTCGACACCACTTCGCCGTGGGCAACATCTT 111

QY 90 CGGGGTGTACCATGTCCAGAACGACTGTCTCCAACTCAAGCATAGTGTATGAGGACGCGGA 149
DB 281 CGCGTGTGGCGCGGACAAAGCTGACCAATCACCAGCGCGGAGCAGGTGGCTGCCGG 340

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCTCGGTTCGGAGGGCAACTCTCTCCGTTG 209
DB 341 CATGGGCATCTAGCGTCCCGCAGCCGCTGTTCTGCAATTGCGCTGAAGGACGCCCGCGCTG 400

QY 210 C 210
DB 401 C 401

RESULT 15
AV632765
LOCUS
DEFINITION
AV632765 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
cDNA clone HC012c10_r 5', mRNA sequence.
ACCESSION
AV632765
VERSION
AV632765.1 GI:10776085
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 473)
AUTHORS
Asanizu,E., Miura,K., Kuchro,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL
DNA Res. 7 (5), 305-307 (2000)
MEDLINE
20539644
PubMed
11089912
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
1..473
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/db_strain="C9"
/db_xref="taxon:3055"
/clone="HC012c10_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      82 a 160 c 151 g 80 t
ORIGIN

Query Match      8.4%; Score 40.2; DB 9; Length 473;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTCTGCTGACCATACAGCTTCCGCTTATGAAGTGGCAACGTGTC 89
DB 156 CTTGACCCCTGGAGCGGTCTGTCATCGTCGACACCACTTCGCCGTGGGCAACATCTT 215

QY 90 CGGGGTGTACCATGTCCAGAACGACTGTCTCCAACTCAAGCATAGTGTATGAGGACGCGGA 149
DB 216 CGCGTGTGGCGCGGACAAAGCTGACCAATCACCAGCGCGGAGCAGGTGGCTGCCGG 275
```


us-09-899-303a-11.rst

Mon Dec 22 13:28:34 2003

QY	150	CATGATCATGCACACCCCGGGTGGTGCCTTCGGGAGGCAACTCCTCCCGTTG	209
DB	276	CATGGGCATCTACGGTCCCGCACCGTGTCTGCATTGCCCTGAAGGACGCCCCGGGCTG	335
QY	210	C	210
DB	336	C	336

Search completed: December 20, 2003, 06:54:43
Job time : 1234.83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 34.6174 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-11

Perfect score: 480

Sequence: 1 ATGTCGGTGTCTTCTC.....TGATGATGAACGTGTAATAG 480

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	3	US-08-612-973-11
2	480	100.0	480	3	US-08-927-597-11
3	474.2	98.8	483	3	US-08-612-973-9
4	474.2	98.8	483	3	US-08-927-597-9
5	431.2	89.8	1539	2	US-08-470-426B-17
6	431.2	89.8	1863	2	US-08-470-426B-14
7	428	89.2	9595	3	US-09-014-416-4
8	428	89.2	9599	3	US-09-014-416-6
9	426.2	88.8	633	3	US-08-612-973-7
10	426.2	88.8	633	3	US-08-927-597-7
11	423.4	88.2	501	2	US-08-483-695-30
12	423.4	88.2	501	2	US-07-965-285-30
13	423.4	88.2	501	2	US-08-487-231-30
14	423.4	88.2	501	3	US-09-010-512-30
15	422.8	88.1	9472	4	US-08-150-204B-96
16	422.2	88.0	642	3	US-08-612-973-3
17	422.2	88.0	642	3	US-08-927-597-3
18	421.2	87.8	795	3	US-08-612-973-5
19	421.2	87.8	795	3	US-08-927-597-5
20	421.2	87.8	2082	3	US-08-612-973-47
21	421.2	87.8	2082	3	US-08-927-597-47
22	421.2	87.8	2433	3	US-08-612-973-49
23	421.2	87.8	2433	3	US-08-927-597-49
24	419.6	87.4	1037	1	US-08-462-195-1
25	419.6	87.4	1037	2	US-08-636-883-1
26	419.6	87.4	1037	3	US-09-127-829-1
27	417.2	86.9	636	3	US-08-612-973-13

28	417.2	86.9	636	3	US-08-927-597-13	Sequence 13, Appl
29	416.8	86.8	742	1	US-08-081-072-18	Sequence 18, Appl
30	416.8	86.8	742	1	US-08-449-093A-18	Sequence 18, Appl
31	416.8	86.8	932	1	US-08-081-072-15	Sequence 15, Appl
32	416.8	86.8	932	1	US-08-449-093A-15	Sequence 15, Appl
33	415.4	86.5	2116	3	US-08-191-160-21	Sequence 21, Appl
34	413.8	86.2	501	2	US-08-483-695-28	Sequence 28, Appl
35	413.8	86.2	501	2	US-07-965-285-28	Sequence 28, Appl
36	413.8	86.2	501	2	US-08-487-231-28	Sequence 28, Appl
37	413.8	86.2	501	3	US-09-201-912-28	Sequence 28, Appl
38	405.2	84.4	1167	1	US-08-324-977-9	Sequence 9, Appl
39	405.2	84.4	1167	2	US-08-384-616-9	Sequence 9, Appl
40	405.2	84.4	1167	2	US-08-904-686A-9	Sequence 9, Appl
41	405.2	84.4	1167	3	US-09-315-850-9	Sequence 9, Appl
42	405.2	84.4	1499	1	US-08-324-977-3	Sequence 3, Appl
43	405.2	84.4	1499	2	US-08-384-616-3	Sequence 3, Appl
44	405.2	84.4	1499	2	US-08-904-686A-3	Sequence 3, Appl
45	405.2	84.4	1499	3	US-09-315-850-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-11
; Sequence 11, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..477
NAME/KEY: mat_peptide
LOCATION: 1..474


```
US-08-612-973-11
Query Match      100.0%; Score 480; DB 3; Length 480;
Best Local Similarity 100.0%; Pred. No. 7.1e-126; Indels 0; Gaps 0;
Matches 480; Conservative 0; Mismatches 0;

;
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..477
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..474
;
; US-08-927-597-11
Query Match      100.0%; Score 480; DB 3; Length 480;
Best Local Similarity 100.0%; Pred. No. 7.1e-126; Indels 0; Gaps 0;
Matches 480; Conservative 0; Mismatches 0;

;
; US-08-612-973-9
; Sequence 9, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
```


CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..480
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..477
US-08-612-973-9

Query Match 98.8%; Score 474.2; DB 3; Length 483;

Best Local Similarity 99.4%; Pred. No. 3e-124;

Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGTCGGTGGCTCTTCTCTATCTTCTTGGCCCTGCTGCTGCTGACCATACCA 60
Db 1 ATGTCGGTGGCTCTTCTCTATCTTCTTGGCCCTGCTGCTGCTGACCATACCA 60
QY 61 GCTTCGGCTTATGAGTGGCGAACGCTGTCGGGGGTGTACCATGTACGACGACTGCTCC 120
Db 61 GCTTCGGCTTATGAGTGGCGAACGCTGTCGGGGGTGTACCATGTACGACGACTGCTCC 120
QY 121 AACTCAAGCATAGTGTATGAGGCGGACATGATCATGCACACCCCGGGTGGCGCC 180
Db 121 AACTCAAGCATAGTGTATGAGGCGGACATGATCATGCACACCCCGGGTGGCGCC 180
QY 181 TCGGTTCCGGAGGGCAACTCCTCCCGTGTGCGGGGTGACATGTCACGACGACTGCTCC 240
Db 181 TCGGTTCCGGAGGGCAACTCCTCCCGTGTGCGGGGTGACATGTCACGACGACTGCTCC 240
QY 241 AGGAACGCGCAGCTCCCAACAGCAATACGACGCGCATGATGCTGCTGCGGGCT 300
Db 241 AGGAACGCGCAGCTCCCAACAGCAATACGACGCGCATGATGCTGCTGCGGGCT 300
QY 301 GCTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360
Db 301 GCTGCTTTCTGTTCCGCTATGATGAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360
QY 361 CAGCTGTTCACTTCTACCTGCGCGGATCAAAAGTACAGACTGCACTGCTCAATC 420
Db 361 CAGCTGTTCACTTCTACCTGCGCGGATCAAAAGTACAGACTGCACTGCTCAATC 420
QY 421 TATCCCGGCCATGATCAGTCACTCCGCTGGGATGATGATGATGATGATGATGATGATGAT 479
Db 421 TATCCCGGCCATGATCAGTCACTCCGCTGGGATGATGATGATGATGATGATGATGATGAT 479

RESULT 4

US-08-927-597-9
Sequence 9, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..480
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..477
US-08-927-597-9

Query Match 98.8%; Score 474.2; DB 3; Length 483;

Best Local Similarity 99.4%; Pred. No. 3e-124;

Matches 476; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGTCGGTGGCTCTTCTCTATCTTCTTGGCCCTGCTGCTGCTGACCATACCA 60
Db 1 ATGTCGGTGGCTCTTCTCTATCTTCTTGGCCCTGCTGCTGCTGACCATACCA 60
QY 61 GCTTCGGCTTATGAGTGGCGAACGCTGTCGGGGGTGTACCATGTACGACGACTGCTCC 120
Db 61 GCTTCGGCTTATGAGTGGCGAACGCTGTCGGGGGTGTACCATGTACGACGACTGCTCC 120
QY 121 AACTCAAGCATAGTGTATGAGGCGGACATGATCATGCACACCCCGGGTGGCGCC 180
Db 121 AACTCAAGCATAGTGTATGAGGCGGACATGATCATGCACACCCCGGGTGGCGCC 180
QY 181 TCGGTTCCGGAGGGCAACTCCTCCCGTGTGCGGGGTGACATGTCACGACGACTGCTCC 240

Db 181 TCGGTTGGAGGGCAATCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGC 240
Qy 241 AGGAAGCCAGCGTCCCCCAACAGCAATACAGCGCCAGTGCATTTGCTGTTGGGGCT 300
Db 241 AGGAAGCCAGCGTCCCCCAACAGCAATACAGCGCCAGTGCATTTGCTGTTGGGGCT 300
Qy 301 GCTGCTTCTGTTCCGCTATGACGTGGGGGATCTCTCGGATCTGTTTCTGTTTCC 360
Db 301 GCTGCTTCTGTTCCGCTATGACGTGGGGGATCTCTCGGATCTGTTTCTGTTTCC 360
Qy 361 CAGCTGTTCACTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATC 420
Db 361 CAGCTGTTCACTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATC 420
Qy 421 TATCCCGGCATGATCAGGTACCGCATCGGCTTGGGATATGATGAATGTTGTAATA 479
Db 421 TATCCCGGCATGATCAGGTACCGCATCGGCTTGGGATATGATGAATGTTGTTCTTA 479

RESULT 5
US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-470-426B-17

Query Match 89.8%; Score 431.2; DB 2; Length 1539;
Best Local Similarity 94.1%; Pred. No. 5.1e-112;
Matches 448; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 2 TGTCCGGTTCCTTTCTCTATCTCTTGGCTTGTCTCTGCTTGTGACCATCCCGAG 565

Db 506 TGCCCGGTTGCTTTTCTATCTATCTTCTTGGCTTGTCTCTGCTTGTGACCATCCCGAG 565
Qy 62 CTTTCCGCTTATGAAGTGCACACGTTGTCGGGGTGTACCATGTACAGAACGACTGCTCCA 121
Db 566 CTTTCCGCTTATGAAGTGCACACGTTGTCGGGGATATACCATGTACAGAACGACTGCTCCA 625
Qy 122 ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGACATCTCCCGGGTGTGCTCCCT 181
Db 626 ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGACATCTCCCGGGTGTGCTCCCT 685
Qy 182 GCGTTCCGGAGGCAACTCTCTCCGTTGCTGGGTGGCTCACTCCACGCTCGCGGCCA 241
Db 686 GCGTTCCGGAGGCAACTCTCTCCGTTGCTGGGTGGCTCACTCCACGCTCGCGGCCA 745
Qy 242 GGAACGCGCGGTCCTCCACAAACGACAAATACGACGCCACGTCGATTTGCTGTTGGGGCTG 301
Db 746 GGAATGCGCGGTCCTCCACAAATACGACAAATACGACGCCACGTCGATTTGCTGTTGGGGCTG 805
Qy 302 CTGCTTTCTGTTCCGCTATGATGAGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
Db 806 CTGCTTTCTGTTCCGCTATGATGAGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 865
Qy 362 AGCTGTTCACTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db 866 AGCTGTTCACTTCTCACTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 925
Qy 422 ATCCCGGCATGATCAGGTACCGCATCGGCTTGGGATATGATGAATGTTGTAATA 477
Db 926 ATCCCGGCATGATCAGGTACCGCATCGGCTTGGGATATGATGAATGTTGTTCA 981

RESULT 6
US-08-470-426B-14
; Sequence 14, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1563 base pairs
; TYPE: nucleic acid


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..627
; US-08-927-597-7

Query Match      88.8%; Score 426.2; DB 3; Length 633;
Best Local Similarity 93.1%; Pred. No. 1e-110;
Matches 446; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 2 TGTCCGGTGTCTTCTCTATCTTCTCTTGTGCGCCCTGCTGCTGTGACCATACCAG 61
Db |||
QY 155 TGCCTGTGTGCTCTTCTCTATCTTCTCTTGTGCTTACTGTCTCTGACCATCCAG 214
Db |||
QY 62 CTTCCGCTTATGAAGTGGCGCAACGTCTCGGGGTGACCATGTGCAGAACGACTGTCCA 121
Db |||
QY 215 CTTCCGCTTATGAAGTGGCGCAACGTCTCGGGGTGACCATGTGCAGAACGACTGTCCA 274
Db |||
QY 122 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGTGCGTCCCT 181
Db |||
QY 275 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGTGCGTCCCT 334
Db |||
QY 182 GGTTCGGAGGGCAACTCTCTCCGTTGCTGGGTGCGTCACTCCACGCTCGCGGCA 241
Db |||
QY 335 GCGTTCGGAGAACAACTCTTCCGCTGTGGGTAGCGCTCACCCGCCACGCTCGCAGTA 394
Db |||
QY 242 GGAACGCGAGCGTCCCGCAACAGACATAGCAGCGCCACGTCGATTTGCTGGGGCTG 301
Db |||
QY 395 GGAACGCGAGCGTCCCGCAACAGACATAGCAGCGCCACGTCGATTTGCTGGGGCTG 454
Db |||
QY 302 CTGCTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTCCTCC 361
Db |||
QY 455 CTGCTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTCCTCC 514
Db |||
QY 362 AGCTGTTCACCTTCTCACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db |||
QY 515 AGCTGTTCACCTTCTCACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 574
Db |||
QY 422 ATCCCGGCGCATGTATCAGTCCAGTCACCGCATGCTTGGGATATGATGAACTGGTAATAG 480
Db |||
QY 575 ATCCCGGCGCATGTATCAGTCCAGTCACCGCATGCTTGGGATATGATGAACTGGTAATAG 633
Db |||

RESULT 11
US-08-483-695-30
; Sequence 30, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brecht, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 08-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; DESCRIPTION: cDNA to genomic RNA
; US-08-483-695-30

Query Match      88.2%; Score 423.4; DB 2; Length 501;
Best Local Similarity 94.4%; Pred. No. 5.8e-110;
Matches 439; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 2 TGTCCGGTGTCTTCTCTATCTTCTCTTGTGCGCCCTGCTGCTGTGACCATACCAG 61
Db |||
QY 37 TGCCTGTGTGCTCTTCTCTATCTTCTCTTGTGCTTGTGCTCTGTTGACCATCCAG 96
Db |||
QY 62 CTTCCGCTTATGAAGTGGCGCAACGTCTCGGGGTGACCATGTGCAGAACGACTGTCCA 121
Db |||
QY 97 CTTCCGCTTATGAAGTGGCGCAACGTCTCGGGGTGACCATGTGCAGAACGACTGTCCA 156
Db |||
QY 122 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGTGCGTCCCT 181
Db |||
QY 157 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCATACTCCCGGTGCGTCCCT 216
Db |||
QY 182 GCGTTCGGAGGGCAACTCTCTCCGTTGCTGGGTGCGCTCACTCCACGCTCGCGGCA 241
Db |||
QY 217 GCGTTCGGAGGGCAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCACGCTCGCGGCA 276
Db |||
QY 242 GGAACGCGAGCGTCCCGCAACAGACATAGCAGCGCCACGTCGATTTGCTGGGGCTG 301
Db |||
QY 277 GGAATGCCAGCGTCCCGCAACAGACATAGCAGCGCCACGTCGATTTGCTGGGGCTG 336
Db |||
QY 302 CTGCTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTCCTCC 361
Db |||
QY 337 CTGCTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTCCTCC 396
Db |||
QY 362 AGCTGTTCACCTTCTCACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db |||
QY 397 AGCTGTTCACCTTCTCACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 456
Db |||
QY 422 ATCCCGGCGCATGTATCAGTCCAGTCACCGCATGCTTGGGATATGATGA 466
Db |||
QY 457 ATCCCGGCGCATGTATCAGTCCAGTCACCGCATGCTTGGGATATGATGA 501
Db |||

RESULT 12
US-07-965-285-30
; Sequence 30, Application US/07965285
; Patent No. 5879904
; GENERAL INFORMATION:

```


Mon Dec 22 13:28:32 2003

APPLICATION NUMBER: KR 91-13601
FILING DATE: 6-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Shahan Islam, Esq.
REGISTRATION NUMBER: 32,507
REFERENCE/DOCKET NUMBER: 2895/F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 940-8564
TELEFAX: (212) 940-8776
INFORMATION FOR SEQ ID NO: 96
SEQUENCE CHARACTERISTICS:
LENGTH: 9472 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: KHCV-LBCL,
SEQUENCE DESCRIPTION: SEQ ID NO: 96

Query Match	88.1%	Score 422.8	DB 4	Length 9472
Best Local Similarity	93.2%	Pred. No. 2e-109	Indels 0	Gaps 0
Matches 442	Conservative 0	Mismatches 32		
QY	2	TGTCGGGTTCGCTCTTTCTCTATCTTCTCTCTTTGGCCCTGCTGCTCTCTGTCGACCATACCAAG	61	
Db	848	TGCCCGGTTCGCTCTTTCTCTATCTTCTCTCTTTGGCTGCTGCTCTGTTGTGGACCAACCCAG	907	
QY	62	CTTCCGCTTATGAAGTCGCGAAAGTGTCGGGGTGACCATGTACAGACGACTGCTCCA	121	
Db	908	TTTCCGCTTATGAAGTCGGTAAACGCTTCGGGGATGTACCATGTACAGAAAGACTGCTCCA	967	
QY	122	ACTCAAGCATAGTGTATGAGSCAGCGACATGATCATGCACACCCCGGGTGGCTGCCCT	181	
Db	968	ACTCAAGCATTTGTGTATGAGSCAGCGACATGATCATGCACACTCCCGGGTGGCTGCCCT	1027	
QY	182	CGCTTCGGGAGGGAACTCTCCCGTTGCTGGGTGGGCTCACTCCACACGCTCGCGGCCA	241	
Db	1028	CGCTTCGGGAGGAACTCTCTCCGTTGCTGGGTGGGCACTTACTCCACACGCTCGCGGCCA	1087	
QY	242	GGACGCCAGCGTCCCAACAACAAATACGACGCCACGTGCGATTTGCTCGTTGGGGCTG	301	
Db	1088	GGATGCCAGCGTCCCACTACGACATTTGCGACGCCATGTGCGACTTGCTGGTGGGTAG	1147	
QY	302	CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGTTTCTCTCTTTCCC	361	
Db	1148	CTGCTTTCTGTTCCGCTATGTACGTGGGGACCTCTCGGATCTGTTTCTCTCTTTCCC	1207	
QY	362	AGCTGTTCACTTCTCACTCTCGCCGGCATCAAAAGTACAGGACTGCAACTGCTCAATCT	421	
Db	1208	AGCTGTTCACTTTTCGCTCTCGCCGGCATGAGCGGTACAGGACTGCAACTGCTCAATCT	1267	
QY	422	ATCCGGCCCATGTATCAGGTACCCGATGGCTTGGGATATGATGATGAACCTGGT	475	
Db	1268	ATCCGGCCCGGTATCAGGTACCGCATGGCTTGGGATATGATGATGAACCTGGT	1321	

Search completed: December 20, 2003, 07:03:01
Job time : 35.6174 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2408.07 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-13
Perfect score: 636
Sequence: 1 ATGCTGGTAGGCGCATCGA.....TGATGACTGGTACTAATAG 636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sta.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sta.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hg_hum.*
- 31: em_hg_inv.*
- 32: em_hg_other.*
- 33: em_hg_mus.*
- 34: em_hg_pln.*
- 35: em_hg_rod.*
- 36: em_hg_mam.*
- 37: em_hg_vrt.*
- 38: em_sy.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	636	100.0	636	6	A48675	A48675 Sequence 13
2	636	100.0	636	6	AR157329	AR157329 Sequence
3	636	100.0	636	6	AX452762	AX452762 Sequence
4	636	100.0	636	6	AX85014	AX85014 Sequence
5	612.8	96.4	633	6	A48669	A48669 Sequence 7
6	612.8	96.4	633	6	AR157326	AR157326 Sequence
7	612.8	96.4	633	6	AX452756	AX452756 Sequence
8	612.8	96.4	633	6	AX685008	AX685008 Sequence
9	604	95.0	795	6	A48667	A48667 Sequence 5
10	604	95.0	795	6	AR157325	AR157325 Sequence
11	604	95.0	795	6	AX452754	AX452754 Sequence
12	604	95.0	795	6	AX685006	AX685006 Sequence
13	601.6	94.6	2082	6	A48709	A48709 Sequence 47
14	601.6	94.6	2082	6	AR157350	AR157350 Sequence
15	601.6	94.6	2082	6	AX452796	AX452796 Sequence
16	601.6	94.6	2082	6	AX85048	AX85048 Sequence
17	601.6	94.6	2433	6	A48711	A48711 Sequence 49
18	601.6	94.6	2433	6	AR157351	AR157351 Sequence
19	601.6	94.6	2433	6	AX452798	AX452798 Sequence
20	601.6	94.6	2433	6	AX685050	AX685050 Sequence
21	566.4	89.1	9379	14	AF207766	AF207766 Hepatitis
22	564.8	88.8	9386	14	AF165056	AF165056 Hepatitis
23	563.2	88.6	9379	14	HCVPOLYP	AJ000009 Hepatitis
24	563.2	88.6	9386	14	AF165055	AF165055 Hepatitis
25	560.6	88.1	9359	14	AF313916	AF313916 Hepatitis
26	560	88.1	9379	14	AF165047	AF165047 Hepatitis
27	560	88.1	9379	14	AF165051	AF165051 Hepatitis
28	560	88.1	9379	14	AF165052	AF165052 Hepatitis
29	560	88.1	9598	14	AB049101	AB049101 Hepatitis
30	558.4	87.8	9379	14	AF165048	AF165048 Hepatitis
31	558.4	87.8	9410	14	HPCCKR2	D50481 Hepatitis C
32	556.8	87.5	1880	14	HPCSTRJ4	D00832 Hepatitis C
33	556.8	87.5	2540	6	E04260	E04260 CDNA encodi
34	556.8	87.5	2540	6	E04805	E04805 CDNA to 5'
35	556.8	87.5	2540	6	E07391	E07391 CDNA encodi
36	556.8	87.5	9344	14	AB049096	AB049096 Hepatitis
37	556.8	87.5	9379	14	AF207761	AF207761 Hepatitis
38	556.8	87.5	9448	14	HPCJ483	D13558 Hepatitis C
39	556.8	87.5	9605	6	AX739971	AX739971 Sequence
40	556.8	87.5	9605	14	HCU238799	AJ238799 Hepatitis
41	556.8	87.5	11076	6	AX036252	AX036252 Sequence
42	556.8	87.5	11076	6	AX036258	AX036258 Sequence
43	556.8	87.5	11076	6	AX036260	AX036260 Sequence
44	556.8	87.5	11076	6	AX036262	AX036262 Sequence
45	556.2	87.5	9431	14	HCU45476	U45476 Hepatitis C

ALIGNMENTS

RESULT 1
A48675
LOCUS A48675
DEFINITION Sequence 13 from Patent WO9604385.
ACCESSION A48675
VERSION A48675.1 GI:2302388
KEYWORDS
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 636)
Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 13 15-FEB-1996;

Mon Dec 22 13:28:34 2003

INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.

FEATURES
source
1..636
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1..633
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAA03133.1"
/db_xref="GI:2302389"
/translation="MLGKAI DLTTCGPADLVGYI PLVGA PLTGGARALAHGVRVLEDD
VNYATNLPGCSFISIFLLALLSCLRIIPASAYEVNVSGMVHTNDCNSSIIVYEAADM
IIMTPGCVPCVRENNSSRCWALVITPLAARNASIFTTIRRHVDLVGAAAFCSAMVY
GDLGSGVFLVSLQFTTISPRRHETVQDNCNSIYPGHITGHRMAWDMNNWY"

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT
ORIGIN

Sequence 13 from patent US 6245503.
AR157329
AR157329.1 GI:16218262
Unknown.
Unidentified.
Unclassified.
1 (bases 1 to 636)
Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M. -A.
Purified hepatitis C virus envelope proteins for diagnostic and
therapeutic use
Patent: US 6245503-A 13 12-JUN-2001;
Location/Qualifiers
1..636
/organism="unknown"
111 a 197 c 175 g 153 t

mat_peptide
1..630
/product="unnamed"
111 a 197 c 175 g 153 t

Query Match
Best Local Similarity
Matches
100.0%; Score 636; DB 6; Length 636;
100.0%; Pred. No. 5.5e-137;
636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAGGCATCGATACCCCTTACGTCGGGCTTCGCCACCTCGTGGGGTACATT 60
Db 1 ATGCTGGGTAAGGCATCGATACCCCTTACGTCGGGCTTCGCCACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCTCCGGGT 120
Db 61 CCGCTCGTCGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCTCCGGGT 120
QY 121 CTGGAAGACGGCGTGAACATATGCAACAGGGAAATTTGGCTGGTCTCTTCTATCTTC 180
Db 121 CTGGAAGACGGCGTGAACATATGCAACAGGGAAATTTGGCTGGTCTCTTCTATCTTC 180
QY 181 CTCCTGGCTTTACTGTCCTGCTTAACCATTCAGCTTCGGCTTACGAGGTGCGCAACGTG 240
Db 181 CTCCTGGCTTTACTGTCCTGCTTAACCATTCAGCTTCGGCTTACGAGGTGCGCAACGTG 240
QY 241 TCCGGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
Db 241 TCCGGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATCATGCAACACCCCGGGTGGTCCCTCGCTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCAACACCCCGGGTGGTCCCTCGCTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGGTAGCGCTCACCCCGGCTAGGAAAGCGGCTAGGAAAGCGGCTAGGAAAGCGG 420
Db 361 TCGTGGGTAGCGCTCACCCCGGCTAGGAAAGCGGCTAGGAAAGCGGCTAGGAAAGCGG 420
QY 421 ATACGACGCGCATCGATTGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCGCATCGATTGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGCGGATCTGCTTCTCTCGTCTCCAGCTGTTCCACATCTCGCTCCCGG 540
Db 481 GGGGATCTCTGCGGATCTGCTTCTCTCGTCTCCAGCTGTTCCACATCTCGCTCCCGG 540
QY 541 CATGAGACGGTGCAGGACTGCAATTCATATCCCGGCGCACATAACGGGTACCGGT 600
Db 541 CATGAGACGGTGCAGGACTGCAATTCATATCCCGGCGCACATAACGGGTACCGGT 600
QY 601 ATGGCTTGGGATATGATGAACTGGTACTAATAG 636
Db 601 ATGGCTTGGGATATGATGAACTGGTACTAATAG 636

RESULT 3
AX452762
LOCUS
DEFINITION
ACCESSION
AX452762
Sequence 13 from Patent EP1211315.
AX452762


```

VERSION      AX452762.1  GI:21712447
KEYWORDS     Hepatitis C virus
SOURCE       Hepatitis C virus
ORGANISM     Hepacivirus.

REFERENCE    1
AUTHORS      Maertens, G., Bosman, F., de Martynoff, G. and Buysse, M.A.
TITLE        Recombinant vectors for producing hcv envelope proteins
JOURNAL      Patent: EP 121315-A 13 05-JUN-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
source       1..636
              /organism="Hepatitis C virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:11103"
              1..633
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAD38071.1"
              /db_xref="GI:21712448"
              /translation="MLGKAIDTLTCGFADLVGYIPLVGAPLGGAAALAHGVRVLEDD
              VNYATGNLPGCSFSLFLALLSLTIPASAYEVNVMYHVTNDNSNIIYVEADNM
              IMHTPCVPCVBNSSRCWALPTPLAARNASIPPTTIRRHVDLLVGGAAAFCSAMYV
              GDLCSVFLVSQLFTISPRRHETVQDCNCSIYPGHITGHRMAMMMWNY"
mat_peptide  1..630
BASE COUNT   111 a 197 c 175 g 153 t
ORIGIN
Query Match      100.0%; Score 636; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 5.5e-137;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGGGGTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGGGGTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTCGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGGTCGGGTT 120
DB 61 CCGCTCGTCGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGGTCGGGTT 120

QY 121 CTGGAAGACGGGTGAACATGCAACAGGGAATTCCTGCTGCTTCTCTATCTTC 180
DB 121 CTGGAAGACGGGTGAACATGCAACAGGGAATTCCTGCTGCTTCTCTATCTTC 180

QY 181 CTCCTGGCTTTACTGCTGTCTTAACCATTTCCAGCTTCCGCTTACGAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTACTGCTGTCTTAACCATTTCCAGCTTCCGCTTACGAGTGGCAACGTG 240

QY 241 TCGGGATGTACCATGTCAAGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGCG 300
DB 241 TCGGGATGTACCATGTCAAGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGCG 300

QY 301 GACATGATCATGCACACCCCGGGTGCCTTCCAGTGTTCACCAATCGCCTCCCGG 360
DB 301 GACATGATCATGCACACCCCGGGTGCCTTCCAGTGTTCACCAATCGCCTCCCGG 360

QY 361 TCGTGGGTAGCGCTACACCCCGGCTCGGGGTAGGAAGCCAGCATCCCCACTACAAC 420
DB 361 TCGTGGGTAGCGCTACACCCCGGCTCGGGGTAGGAAGCCAGCATCCCCACTACAAC 420

QY 421 ATACGACGCCACATCGATTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
DB 421 ATACGACGCCACATCGATTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480

QY 481 GGGGATCTCTGCGGATCTGCTTCTGCTTCCAGTGTTCACCAATCGCCTCCCGG 540
DB 481 GGGGATCTCTGCGGATCTGCTTCTGCTTCCAGTGTTCACCAATCGCCTCCCGG 540

QY 541 CATGACGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAACGGGTACCGT 600
DB 541 CATGACGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAACGGGTACCGT 600

QY 601 ATGGCTTGGGATATGATGATGATGAAGTGGTACTAATAG 636
DB 601 ATGGCTTGGGATATGATGATGATGAAGTGGTACTAATAG 636

RESULT 4
AX452762.1  GI:21712447
LOCUS       AX452762.1  636 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 13 from Patent WO0205548.
ACCESSION   AX452762.1
VERSION     AX452762.1  GI:21712447
KEYWORDS    Hepatitis C virus
SOURCE      Hepatitis C virus
ORGANISM    Hepacivirus.
REFERENCE   1
AUTHORS     Maertens, G., Bosman, F. and Buysse, M.A.
TITLE       Purified Hepatitis C Virus envelope proteins for diagnostic and
            therapeutic use
JOURNAL     Patent: WO 0205548-A 13 18-JUL-2002;
            INNOGENETICS N.V. (BE)
FEATURES    Location/Qualifiers
source      1..636
            /organism="Hepatitis C virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:11103"
            1..633
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD86524.1"
            /db_xref="GI:29371420"
            /translation="MLGKAIDTLTCGFADLVGYIPLVGAPLGGAAALAHGVRVLEDD
            VNYATGNLPGCSFSLFLALLSLTIPASAYEVNVMYHVTNDNSNIIYVEADNM
            IMHTPCVPCVBNSSRCWALPTPLAARNASIPPTTIRRHVDLLVGGAAAFCSAMYV
            GDLCSVFLVSQLFTISPRRHETVQDCNCSIYPGHITGHRMAMMMWNY"
mat_peptide 1..630
BASE COUNT  111 a 197 c 175 g 153 t
ORIGIN
Query Match      100.0%; Score 636; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 5.5e-137;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGGGGTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGGGGTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTCGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGGTCGGGTT 120
DB 61 CCGCTCGTCGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGGTCGGGTT 120

QY 121 CTGGAAGACGGGTGAACATGCAACAGGGAATTCCTGCTGCTTCTCTATCTTC 180
DB 121 CTGGAAGACGGGTGAACATGCAACAGGGAATTCCTGCTGCTTCTCTATCTTC 180

QY 181 CTCCTGGCTTTACTGCTGTCTTAACCATTTCCAGCTTCCGCTTACGAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTACTGCTGTCTTAACCATTTCCAGCTTCCGCTTACGAGTGGCAACGTG 240

QY 241 TCGGGATGTACCATGTCAAGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGCG 300
DB 241 TCGGGATGTACCATGTCAAGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGCG 300

QY 301 GACATGATCATGCACACCCCGGGTGCCTTCCAGTGTTCACCAATCGCCTCCCGG 360
DB 301 GACATGATCATGCACACCCCGGGTGCCTTCCAGTGTTCACCAATCGCCTCCCGG 360

QY 361 TCGTGGGTAGCGCTACACCCCGGCTCGGGGTAGGAAGCCAGCATCCCCACTACAAC 420
DB 361 TCGTGGGTAGCGCTACACCCCGGCTCGGGGTAGGAAGCCAGCATCCCCACTACAAC 420

QY 421 ATACGACGCCACATCGATTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
DB 421 ATACGACGCCACATCGATTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480

QY 481 GGGGATCTCTGCGGATCTGCTTCTGCTTCCAGTGTTCACCAATCGCCTCCCGG 540
DB 481 GGGGATCTCTGCGGATCTGCTTCTGCTTCCAGTGTTCACCAATCGCCTCCCGG 540

QY 541 CATGACGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAACGGGTACCGT 600
DB 541 CATGACGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAACGGGTACCGT 600
```


421	ATACAGCCACGTCGATTTGCTTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATATACGTTG	480
421	ATACAGCCACGTCGATTTGCTTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATATACGTTG	480
481	GGGGATCTCTGCGGATCTGTCTTCCTGTTCTCCAGCTGTTCCACCATCTCGCTCGCGG	540
481	GGGGATCTCTGCGGATCTGTCTTCCTCGTCTCCAGCTGTTCCACCATCTCGCTCGCGG	540
541	CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGGCCACATAACCGGTCAACGTT	600
541	CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGGCCACATAACCGGTCAACGTT	600
601	ATGGCTTGGGATATGATGATGAACCTGGTACTAATAG	636
601	ATGGCTTGGGATATGATGATGAACCTGGTACTAATAG	636

RESULT 5
 A48669
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source
 CDS
 linear PAT 07-MAR-1997
 633 bp DNA
 Sequence 7 from Patent WO9604385.
 A48669
 A48669.1 GI:2302382
 unidentified
 unclassified
 1 (bases 1 to 633)
 Maertens,G., Bosman,F., De,M.G. and Buysse,M.
 PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
 THERAPEUTIC USE
 Patent: WO 9604385-A 7 15-FEB-1996;
 INNOGENETICS NV (BE)
 Other publication CA 2172273 960215
 Other publication AU 3382495 960304.
 Location/Qualifiers
 1..633
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 1..630
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAA03130.1"
 /db_xref="GI:2302383"
 /translation="MLGKVIDTLTCGFADLMGVIPLVGAPLGGARALAHGVRVLVED
 NYIAIGNLPFGCSFSIFLLALSLCTIPASAVEYRNVMGVHTVNDCSNSIVYEAADM
 IMHTPGCVPCVRENNSSRCWTALPTLAARNAASVPTTIRRHVDLLVGAFAACSMYV
 "

Authors Title	Journal	Comment	Features	Source	CDS	mat_peptide	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches 620;	Conservative	Indels	Gaps	Length 633;
PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE	PATENT: WO 9604385-A 7 15-FEB-1996;	INNOGENETICS NV (BE) Other publication CA 2172273 960215 Other publication AU 3382495 960304.	Location/Qualifiers	1. .633 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"	1. .630 /note="unnamed protein product"	1. .627 /product="unnamed"	111 a 192 c 174 g 156 t	86.4%; Score 612.8; DB 6; 98.1%; Pred. No. 1.3e-13; 0; Mismatches 12;							

mat_peptide	1..627	/product="unnamed"	111 a	192 c	174 g	156 t	
BASE COUNT							
ORIGIN							
Query Match		96.4%;	Score 612.8;	DB 6;	Length 633;		
Best Local Similarity		98.1%;	Pred. No. 1.3e-13;				
Matches 620;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;		
Qy	1	ATGCTGGGTAAAGCCATCGATACCCCTTACGTGCGGCTTGC	CGACCTCGTGGGGTACATT	60			
Db	1	ATGTTGGGTAAAGTCATCGATACCCCTTACGTGCGGCTTGC	CGACCTCATGGGGTACATT	60			
Qy	61	CCGCTCGTGGGGCCCCCTTAGGGGGCGTCGACGGGCCCTGG	CGCAGTGGCTCCGGGT	120			
Db	61	CCGCTCGTGGGGCCCCCTTAGGGGGTGTGCGAGAGCCCTGG	CGCATGGCTCCGGGT	120			
Qy	121	CTGGAGACGGCGGTGAACATATGCAACAGGGAAATTGCC	TGGTTGCTTTCTATCTTC	180			
Db	121	CTGGAGACGGCGGTGAACATATGCAACAGGGAAATTGCC	TGGTTGCTTTCTATCTTC	180			
Qy	181	CTCTTGGCGTTTACTGTCCTGCTTAAACATTTCAGCTTCG	CTTACGAGTGGCGCACGTTG	240			


```
Db 241 TCCTGGATGATACATGTCAGAACGACTGCTCAACTCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGTTTCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGTTTCGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGGTAGCGTCAACCCCGAGCTCGCGGTAGGAACGCCAGCATCCCACTACAACA 420
Db 361 TCGTGGGTAGCGTCAACCCCGAGCTCGAGCTAGGAACGCCAGCTCCCACTAGACA 420
QY 421 ATACGACGCCACGTCGATTTGCTGCTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGTCGATTTGCTGCTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGGGATCTGCTTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
Db 481 GGGGATCTCTGGGATCTGCTTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
QY 541 CATGACAGCGTGCAGACGTCGATTTGCTGCTTCCAGCTGTTCAACCATCTCGCCTCGCGG 600
Db 541 CATGACAGCGTGCAGACGTCGATTTGCTGCTTCCAGCTGTTCAACCATCTCGCCTCGCGG 600
QY 601 ATGGCTTGGGATGATGATGAACGTTACTA 632
Db 601 ATGGCTTGGGATGATGATGAACGTTACTA 632

RESULT 7
AX452756
LOCUS AX452756 633 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 7 from Patent EP1211315.
ACCESSION AX452756
VERSION AX452756.1 GI:21712441
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; SERNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buysse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 7 05-JUN-2002;
INNOCENETICS N.V. (BE)
FEATURES
source
1..633
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..630
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD38068.1"
/db_xref="GI:21712442"
VNYATNLPKCSFISPLALLSLTIPASAYEVNVSVMYHVTNDCNSISIVYEAADM
INHTPCVPCVBENSSRCWALTPTLAARASVPTTIRRHVDLLVGAARAFCSAMYV
GDLCSGVFLVSQFLTISPRHETVQDCNCSIIYPGHITGHRMAMMMNNW"
mat_peptide 1..627
/product="unnamed"
BASE COUNT 111 a 192 c 174 g 156 t
ORIGIN
Query Match 96.4%; Score 612.8; DB 6; Length 633;
Best Local Similarity 98.1%; Pred. No. 1.3e-131;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATCTGGGTAAAGCCATCGATACCTTACGTGCGGTTCCGCCACCTCGTGGGTACATT 60
Db 1 ATCTGGGTAAAGTCATCGATACCTTACGTGCGGTTCCGCCACCTCATGGGTACATT 60
QY 61 CCCTCTGTCGGCCCCCTTAGGGGGCGTTCGACGGGCCCTGCGCATGCGTCCGGTT 120
```

```
Db 61 CCGCTCTGTCGGCGCCCCCTAGGGGTGCTGCCAGAGCCCTGGCGCATGGCGTCCGGTT 120
QY 121 CTGGAAGACGGGTGAATCTATGCAACAGAGGAATTTGCTGTTGCTTCTTCTATCTTC 180
Db 121 CTGGAAGACGGGTGAATCTATGCAACAGAGGAATTTGCTGTTGCTTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTACTGCTCTTAACCATTCAGCTTCCGCTTACAGGTGGCAACGTTG 240
Db 181 CTCTTGGCTTTACTGCTCTTAACCATTCAGCTTCCGCTTATGAGGTGGCAACGTTG 240
QY 241 TCGGGATGTCACATGTCACGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGAGCG 300
Db 241 TCGGGATGTCACATGTCACGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGTTTCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGTTTCGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGGTAGCGTCAACCCCGAGCTCGCGGTAGGAACGCCAGCATCCCACTACAACA 420
Db 361 TCGTGGGTAGCGTCAACCCCGAGCTCGAGCTAGGAACGCCAGCTCCCACTAGACA 420
QY 421 ATACGACGCCACGTCGATTTGCTGCTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGTCGATTTGCTGCTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGGGATCTGCTTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
Db 481 GGGGATCTCTGGGATCTGCTTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
QY 541 CATGACAGCGTGCAGACGTCGATTTGCTGCTTCCAGCTGTTCAACCATCTCGCCTCGCGG 600
Db 541 CATGACAGCGTGCAGACGTCGATTTGCTGCTTCCAGCTGTTCAACCATCTCGCCTCGCGG 600
QY 601 ATGGCTTGGGATGATGATGAACGTTACTA 632
Db 601 ATGGCTTGGGATGATGATGAACGTTACTA 632

RESULT 8
AX685008
LOCUS AX685008 633 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 7 from Patent WO02055548.
ACCESSION AX685008
VERSION AX685008.1 GI:29371413
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; SERNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F. and Buysse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: WO 02055548-A 7 18-JUL-2002;
INNOCENETICS N.V. (BE)
FEATURES
source
1..633
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..630
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD86521.1"
/db_xref="GI:29371414"
VNYATNLPKCSFISPLALLSLTIPASAYEVNVSVMYHVTNDCNSISIVYEAADM
INHTPCVPCVBENSSRCWALTPTLAARASVPTTIRRHVDLLVGAARAFCSAMYV
GDLCSGVFLVSQFLTISPRHETVQDCNCSIIYPGHITGHRMAMMMNNW"
mat_peptide 1..627
/product="unnamed"
BASE COUNT 111 a 192 c 174 g 156 t
```


ORIGIN		CDS		/db_xref="taxon:32644"	
Query Match		1..792		/note="unnamed protein product"	
Best Local Similarity		95.0%; Score 604; DB 6; Length 795;		/codon_start=1	
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;		97.6%; Pred. No. 1.3e-131; Mismatches 15; Indels 0; Gaps 0;		/protein_id="CAA03129.1"	
1		ATGCTGGTAAAGCCATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		/db_xref="GI:2302381"	
60				/translation="MLGKVIDLTTCGFDLVGYIPLVGAPLGPAGARALAHGVRLVEDG	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		VNYATGNLPGCFPSIFLLALLSLCTVPASAYEVNRVNGMYHTNDCNSNIIYERADM	
60				IHTPCVPCVRENSSRCWALTPTLAARNASVPTTIRRHVDLLVGAATFCASMYV	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QLLGRIFQAVVDVAVAGAHWGLAGLAYYSVMGNWAKVLIIVMLLPAP"	
120				mat_peptide	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		1..799	
60				/product="unnamed"	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		BASE COUNT	
120				130 a 240 c 231 g 194 t	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		ORIGIN	
60				Query Match	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Best Local Similarity	
120				Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				1	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		61	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				61	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		121	
120				Db	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		QY	
60				121	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		Db	
120				QY	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		121	
60				Db	
61		CCGCTCGTGGGCGCCCTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGGT		QY	
120				121	
1		ATGTTGGGTAAAGTCAATCATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT		Db	
60				QY	

QY	481	GGGATCTTCGGGATCTGCTCTCTCCAGCTGTTTCCACCATCTCGCTCGCCGG	540
Db	481	GGGGACCTCTGGGGATCTGCTCTCTCCAGCTGTTTCCACCATCTCGCTCGCCGG	540
QY	541	CATGAGACGGTGCAGGACTGCAATTGCTCAATTCATCCGGCCACATACCGGTACCGT	600
Db	541	CATGAGACGGTGCAGGACTGCAATTGCTCAATTCATCCGGCCACATACCGGTACCGT	600
QY	601	ATGCGTTGGGATATGATGATGAACCTGGT	628
Db	601	ATGCGTTGGGATATGATGATGAACCTGGT	628
RESULT 13			
LOCUS	A48709	2082 bp	DNA
DEFINITION	Sequence 47 from Patent WO9604385.		
ACCESSION	A48709		
VERSION	A48709.1	GI:2302422	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 2082)		
AUTHORS	Maertens, G., Bosman, F., De, M. G. and Buyse, M.		
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE		
JOURNAL	PATENT: WO 9604385-A 47 15-FEB-1996;		
COMMENT	INNOGENETICS NV (BE)		
	Other publication CA 2172273 960215		
	Other publication AU 3382495 960304.		
FEATURES	Location/Qualifiers		
source	1..2082		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
CDS	1..2079		
	/note="unnamed protein product; Protein sequence is in conflict with the conceptual translation"		
	/codon_start=1		
	/protein_id="CAA03146.1"		
	/db_xref="GI:2302423"		
	/translation="NLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRVLEDD VNYATGNLPGCSFSLFLLALLSCLTVPASAYEVNRVSGMVHTNDCNSISIVYEAADM IMHTPGCPVCVRENSSRCWALTPTLAARNASVPTTIRRHVDLLVGAAPFCSAMVY GDLCSFVLVSQFLTISPRRHETVQDCNCISYPGHIITGHRMADMMWNSPTTALVVS QLLRIPOAVDMVAGAHGVLAGLAYSVMGNWAKVLIIVMLLPAP"		
mat_peptide	1..789		
BASE COUNT	130 a 240 c 231 g 194 t		
ORIGIN	/product="unnamed"		
Query Match	95.0%; Score 604; DB 6; Length 795;		
Best Local Similarity	97.6%; Pred. No. 1.4e-129;		
Matches	613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;		
QY	1	ATGCTGGGTAAAGCCATGATACCTTACGTGCGGCTCGCGACCTCGTGGGTACATT	60
Db	1	ATGTTGGGTAAAGTCATGATACCTTACGTGCGGCTTCGCCGACCTCGTGGGTACATT	60
QY	61	CCGCTCGTGGGCGCCCTAGGGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGTT	120
Db	61	CCGCTCGTGGGCGCCCTAGGGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGTT	120
QY	121	CTGGAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTCTATCTTC	180
Db	121	CTGGAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTCTATCTTC	180
QY	181	CTCTTGGCTTACTGTCCTGCTTCACTTCCAGCTTCCGCTTACGAGTGGCAACGTG	240
Db	181	CTCTTGGCTTACTGTCCTGCTTCACTTCCAGCTTCCGCTTATGAGTGGCAACGTG	240
QY	241	TCCGGATGATACCATGTACGACGACTGCTCAACTCAAGCATTTGTATGAGGACGG	300
Db	241	TCCGGATGATACCATGTACGACGACTGCTCAACTCAAGCATTTGTATGAGGACGG	300
QY	301	GACATGATCATGACACCCCGGTTGGTCCCTGCTTCCGGAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGACACCCCGGTTGGTCCCTGCTTCCGGAGAACAACTCTTCCCGC	360
QY	361	TGCTGGGTAGCGTCAACCCACGCTCGCGGTAGGAAACGCGAGCATCCCACTCAACA	420
Db	361	TGCTGGGTAGCGTCAACCCACGCTCGCGGTAGGAAACGCGAGCATCCCACTCAACA	420
QY	421	ATACAGCCGACGTCGATTTGCTCGTTGGGCGCGCTCTTCTTCTTCCGCTATGTACGTG	480
Db	421	ATACAGCCGACGTCGATTTGCTCGTTGGGCGCGCTCTTCTTCTTCCGCTATGTACGTG	480
LOCUS	AX685006	795 bp	DNA
DEFINITION	Sequence 5 from Patent WO02055548.		
ACCESSION	AX685006		
VERSION	AX685006.1	GI:29371411	
KEYWORDS	Hepatitis C virus		
SOURCE	Hepatitis C virus		
ORGANISM	Viruses; SARNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1		
AUTHORS	Maertens, G., Bosman, F. and Buyse, M.A.		
TITLE	Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use		
JOURNAL	Patent: WO 02055548-A 5 18-JUL-2002;		
	INNOGENETICS N.V. (BE)		
FEATURES	Location/Qualifiers		
source	1..795		
	/organism="Hepatitis C virus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:11103"		
CDS	1..792		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD86520.1"		
	/db_xref="GI:29371412"		
	/translation="MUGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRVLEDD VNYATGNLPGCSFSLFLLALLSCLTVPASAYEVNRVSGMVHTNDCNSISIVYEAADM IMHTPGCPVCVRENSSRCWALTPTLAARNASVPTTIRRHVDLLVGAAPFCSAMVY GDLCSFVLVSQFLTISPRRHETVQDCNCISYPGHIITGHRMADMMWNSPTTALVVS QLLRIPOAVDMVAGAHGVLAGLAYSVMGNWAKVLIIVMLLPAP"		
mat_peptide	1..789		
BASE COUNT	130 a 240 c 231 g 194 t		
ORIGIN	/product="unnamed"		
Query Match	95.0%; Score 604; DB 6; Length 795;		
Best Local Similarity	97.6%; Pred. No. 1.4e-129;		
Matches	613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;		
QY	1	ATGCTGGGTAAAGCCATGATACCTTACGTGCGGCTCGCGACCTCGTGGGTACATT	60
Db	1	ATGTTGGGTAAAGTCATGATACCTTACGTGCGGCTTCGCCGACCTCGTGGGTACATT	60
QY	61	CCGCTCGTGGGCGCCCTAGGGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGTT	120
Db	61	CCGCTCGTGGGCGCCCTAGGGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGTT	120
QY	121	CTGGAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTCTATCTTC	180
Db	121	CTGGAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTCTATCTTC	180
QY	181	CTCTTGGCTTACTGTCCTGCTTCACTTCCAGCTTCCGCTTACGAGTGGCAACGTG	240
Db	181	CTCTTGGCTTACTGTCCTGCTTCACTTCCAGCTTCCGCTTATGAGTGGCAACGTG	240
QY	241	TCCGGATGATACCATGTACGACGACTGCTCAACTCAAGCATTTGTATGAGGACGG	300
Db	241	TCCGGATGATACCATGTACGACGACTGCTCAACTCAAGCATTTGTATGAGGACGG	300
QY	301	GACATGATCATGACACCCCGGTTGGTCCCTGCTTCCGGAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGACACCCCGGTTGGTCCCTGCTTCCGGAGAACAACTCTTCCCGC	360
QY	361	TGCTGGGTAGCGTCAACCCACGCTCGCGGTAGGAAACGCGAGCATCCCACTCAACA	420
Db	361	TGCTGGGTAGCGTCAACCCACGCTCGCGGTAGGAAACGCGAGCATCCCACTCAACA	420
QY	421	ATACAGCCGACGTCGATTTGCTCGTTGGGCGCGCTCTTCTTCTTCCGCTATGTACGTG	480
Db	421	ATACAGCCGACGTCGATTTGCTCGTTGGGCGCGCTCTTCTTCTTCCGCTATGTACGTG	480

Mon Dec 22 13:28:34 2003

```
mat_peptide 1..2076
BASE COUNT 366 a 634 c 600 g 482 t
ORIGIN
Query Match 94.6%; Score 601.6; DB 6; Length 2082;
Best Local Similarity 97.8%; Pred. No. 5.3e-129;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 TGGGTAAGGCCATCGATACCTTACGTGGGCTTCGCCGACCTCGTGGGTACATTCGCG 64
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 5 TGGGTAAGGTCATCGATACCTTACATCGGGCTTCGCCGACCTCGTGGGTACATTCGCG 64
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 TCCTCGGCGCCCTTCAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCGG 124
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 TCCTCGGCGCCCTTCAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCGG 124
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 125 AAGACGGCGTGAACATATGCAACAGGGAATTGCTGGTTCCTTTCTATCTCTCTCT 184
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 125 AAGACGGCGTGAACATATGCAACAGGGAATTGCTGGTTCCTTTCTATCTCTCTCT 184
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 185 TGGCTTTATCTGCTCTGTAACTTCCAGCTTCGCTTACGAGTTCGCAACGTGTCCG 244
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 185 TGGCTTTATCTGCTCTGTAACTTCCAGCTTCGCTTACGAGTTCGCAACGTGTCCG 244
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 245 GGATGTACCATGTCCAGAACGACTGTCTCCAACCTCAAGCAATTGTATGAGGCGGACA 304
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 245 GGATGTACCATGTCCAGAACGACTGTCTCCAACCTCAAGCAATTGTATGAGGCGGACA 304
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 305 TGATCATGCAACACCCCGGGTGGTCCCTGCGTTTCGGGAGAACAACTCTTCCCGCTGCT 364
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 305 TGATCATGCAACACCCCGGGTGGTCCCTGCGTTTCGGGAGAACAACTCTTCCCGCTGCT 364
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 365 GGGTAGCGCTCACCCCGGCTCGCGGTAGAACGCGCAGCATCCCCACTTACAAACATAC 424
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 365 GGGTAGCGCTCACCCCGGCTCGCGGTAGAACGCGCAGCATCCCCACTTACAAACATAC 424
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 425 GACGCCAGCTCGAATTGCTCGTTGGGGCGGTGCTTTCTGTTCCGCTATGATGTTGGGG 484
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 425 GACGCCAGCTCGAATTGCTCGTTGGGGCGGTGCTTTCTGTTCCGCTATGATGTTGGGG 484
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 485 ATCTCTGGGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 544
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 485 ATCTCTGGGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 544
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 545 AGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCAATACCGGTCACCGTATGG 604
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 545 AGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCAATACCGGTCACCGTATGG 604
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 605 CTTGGGATATGATGAAGTGGT 628
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 605 CTTGGGATATGATGAAGTGGT 628
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: December 20, 2003, 02:01:55
Job time : 2410.07 secs


```
Db 181 CTTCTGGCTTACTGTCCTGCTTACCAATTCAGCTTCCGTTTACAGGTGCCAACGTTG 240
QY 241 TCCGGGATGATACATGATCAGAACGAGTCTCTCAACTCAAGCAATTTGTATGAGGAGCG 300
Db 241 TCCGGGATGATACATGATCAGAACGAGTCTCTCAACTCAAGCAATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGACACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 GACATGATCATGACACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TGTCTGGTACGCTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 TGTCTGGTACGCTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 ATAGAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 ATAGAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GGGGATCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GGGGATCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CATGAGACGCTGACGAGCTGCAATTTGCTCAATCTATCTATCTATCTATCTATCTATCT 600
Db 541 CATGAGACGCTGACGAGCTGCAATTTGCTCAATCTATCTATCTATCTATCTATCTATCT 600
QY 601 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 636
Db 601 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 636
```

RESULT 3

```
AA12706
ID AAT12706 standard; DNA; 633 BP.
XX
AC AAT12706;
XX
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC111A.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW SB.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
DR WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di: sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins:
CC The recombinant proteins can then be isolated using a method of the
```

```
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
```

SQ Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;

Query Match 96.4%; Score 612.8; DB 17; Length 633;
Best Local Similarity 98.1%; Pred. No. 1.3e-155;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAGGCCATCGATACCCCTTACGTGCGGCTTCGCCACCTCGTGGGTACATT 60
Db 1 ATGCTGGGTAAGGTCATCGATACCCCTTACGTGCGGCTTCGCCACCTCATGGGTACATT 60
QY 61 CGCTCGTCGGGCGCCCTTAGGGGCGCTGCGAGGCGCTGCGCGCATGCGCGTTCGGGTT 120
Db 61 CGCTCGTCGGGCGCCCTTAGGGGCGCTGCGAGGCGCTGCGCGCATGCGCGTTCGGGTT 120
QY 121 CTGGAAGCGGCGTGAATATGCAACAGGGAATTTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 CTGGAAGCGGCGTGAATATGCAACAGGGAATTTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 181 CTCTGGCTTTACTGCTGCTGCTTAACCATTTCCAGCTTCCGCTTACGAGGTGCCAACGTTG 240
Db 181 CTCTGGCTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 TCCGGATGTACATGTACGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 TCCGGATGTACATGTACGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 301 GACATGATCATGACACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 GACATGATCATGACACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 TGTCTGGTACGCTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 TGTCTGGTACGCTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 ATAGAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 ATAGAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GGGGATCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GGGGATCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CATGAGACGCTGACGAGCTGCAATTTGCTCAATCTATCTATCTATCTATCTATCTATCT 600
Db 541 CATGAGACGCTGACGAGCTGCAATTTGCTCAATCTATCTATCTATCTATCTATCTATCT 600
QY 601 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 632
Db 601 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 632

RESULT 4

AA148915
ID AAL48915 standard; DNA; 633 BP.
XX
AC AAL48915;
XX
DT 24-OCT-2002 (first entry)

XX Hepatitis C virus clone HCC111A E1 protein coding sequence.
 DE Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 XX virucide; immunostimulant; vaccine; db.
 XX Hepatitis C virus.
 XX WO200255548-A2.
 XX 18-JUL-2002.
 XX 11-JAN-2002; 2002WO-EP00219.
 XX 11-JAN-2001; 2001US-260699P.
 XX 30-AUG-2001; 2001US-315768P.
 XX (INNO-) INNOGENETICS NV.
 XX Maertens G, Bosman F, Buyse M;
 XX WPI; 2002-599657/64.
 XX P-PSDB; AAO18662.
 XX New therapeutic vaccine compositions comprising at least one purified
 FT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans
 PT from HCV infection -
 XX
 XX Example 2; Page 163-164; 243pp; English.
 XX The present invention relates to new therapeutic vaccine compositions for
 CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosis. The response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.
 XX
 XX Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;
 SQ
 Query Match 96.4%; Score 612.8; DB 24; Length 633;
 Best Local Similarity 98.1%; Pred. No. 1.3e-155;
 Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 1 ATGCTGGGTAAGGCCATCGATACCTTACGTGCGGCTTCGCGACCTCGTGGGTACATT 60
 1 ATGTTGGGTAAGGTATCGATACCTTACGTGCGGCTTCGCGACCTCATGGGTACATT 60
 61 CCGCTCGTGGGCGCCCTAGGGGGCTGCGAGGCGCTGGCGCATGGCGTCCGGGTT 120
 61 CCGCTCGTGGGCGCCCTAGGGGGCTGCGAGGCGCTGGCGCATGGCGTCCGGGTT 120
 121 CTGGAAGCGGCTGAACATATGCAACAGGGAATTTGCCCTGGTTGCTTCTATCTTC 180
 121 CTGGAAGCGGCTGAACATATGCAACAGGGAATTTGCCCTGGTTGCTTCTATCTTC 180
 181 CTCTGGCTTACTGCTGCTGAGCAATTCACATCCAGCTTCAGAGTGGCAACGTG 240
 181 CTCTGGCTTACTGCTGCTGAGCAATTCACATCCAGCTTCAGAGTGGCAACGTG 240
 241 TCCGGATGTACCATGTGACAGCACTGCTCCAACTCAAGCATTTGTGTATGAGCAGCG 300
 241 TCCGGATGTACCATGTGACAGCACTGCTCCAACTCAAGCATTTGTGTATGAGCAGCG 300
 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGTTCGGGAGAACACTTCCCGC 360
 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGTTCGGGAGAACACTTCCCGC 360

QY 361 TGCTGGGTAGCGCTACCCCGAGCTCGCGGCTAGGAAGCCAGCATCCCACTACAACA 420
 Db 361 TGCTGGGTAGCGCTACCCCGAGCTCGCGGCTAGGAAGCCAGCATCCCACTACAACA 420
 QY 421 ATACGAGCGGCTCGATTTGCTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGACGTG 480
 Db 421 ATACGAGCGGCTCGATTTGCTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGACGTG 480
 QY 481 GGGGATCTCTGCGGATCTGTCTTCCTGCTCCAGCTGTTACCATCTCGGCTCGCCGG 540
 Db 481 GGGGATCTCTGCGGATCTGTCTTCCTGCTCCAGCTGTTACCATCTCGGCTCGCCGG 540
 QY 541 CATGAGCGGCTGAGGACTGCAATTCCTCAATCTATCCCGGCCACATAACGGGTACCGT 600
 Db 541 CATGAGCGGCTGAGGACTGCAATTCCTCAATCTATCCCGGCCACATAACGGGTACCGT 600
 QY 601 ATGGCTTGGGATATGATGATGAAGTGAACCTGCTACTA 632
 Db 601 ATGGCTTGGGATATGATGATGAAGTGAACCTGCTAATA 632
 RESULT 5
 AAT12705
 ID AAT12705 standard; DNA; 795 BP.
 XX AAT12705;
 AC AAT12705;
 XX 23-SEP-1996 (first entry)
 DT HCV E1 construct HCC110A.
 DE HCV E1 construct HCC110A.
 XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.
 XX Hepatitis C virus.
 XX WO9604385-A2.
 XX 15-FEB-1996.
 XX 31-JUL-1995; 95WO-EP03031.
 XX 29-JUL-1994; 94EP-0870132.
 XX (INNO-) INNOGENETICS NV.
 XX Bosman F, Buyse M, De Martynoff G, Maertens G;
 WPI; 1996-129401/13.
 Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 proteins - in presence of disulphide bond cleavage agent, to
 produce proteins suitable for direct use in vaccines or diagnostic
 assays of HCV
 Claim 23; Fig 21; 146pp; English.
 AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 and E2 protein coding sequence constructs. These sequences are included
 in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 The recombinant proteins can then be isolated using a method of the
 invention. In the method, the envelope proteins are purified by
 carrying out a disulphide bond cleavage, or a reduction step with a
 disulphide bond cleavage agent, after lysis of recombinant host cells.
 The constructs containing the purified HCV envelope proteins can be used
 for vaccinating humans against HCV, for in vitro detection of HCV
 antibodies in a sample, and in a serotyping assay for detecting one or
 more serological types of HCV present in a biological sample. The
 constructs can also be immobilised on a solid substrate and incorporated
 into a reversed phase hybridisation assay for determining the presence or
 the genotype of HCV. The new purification method preserves the

CC	conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.
XX	
SQ	Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;
	Query Match 95.0%; Score 604; DB 17; Length 795;
	Best Local Similarity 97.6%; Pred. No. 3.3e-153;
	Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY	1 ATGCTGGTAAAGGCATCGATACCCCTTACGTCGGCTTCGCCGACCTCGTGGGTACATT 60
DB	1 ATGTTGGTAAGGTATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY	61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB	61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
QY	121 CTGGAAGACGGCGTGAATATGCAACAGGGAATTTGGCTGGTCTCTTCTCTATCTTC 180
DB	121 CTGAGAGACGGCGTGAATATGCAACAGGGAATTTGGCCGGTTGCTCTTCTCTATCTTC 180
QY	181 CTCTTGGCTTTACTGTCCTGCTTAACCATTCACAGCTTCGGCTTACGAGTGGCAAGCTG 240
DB	181 CTCTTGGCTTTGCTGTCTGCTGTACCGTTCCAGCTTCGGCTTATGAAGTGGCAAGCTG 240
QY	241 TCCGGGATGTACCATGTACAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCAGCG 300
DB	241 TCCGGGATGTACCATGTACAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCAGCG 300
QY	301 GACATGATCATGCACACCCCGGGTGGCTGCCCTTCCGCTTCGGGAGAACAACTCTTCCCGC 360
DB	301 GACATGATCATGCACACCCCGGGTGGCTGCCCTTCCGCTTCGGGAGAACAACTCTTCCCGC 360
QY	361 TGCTGGGTAGCGCTACACCCACGCTCGCGGTAGGACGCGCAGCATCCCACTACAAACA 420
DB	361 TGCTGGGTAGCGCTACACCCACGCTCGCAGCTAGGAAACGCGCAGCTCCCAACCAACACA 420
QY	421 ATACGACGCCACGTCGATTTGCTGGTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTG 480
DB	421 ATACGACGCCACGTCGATTTGCTGGTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTG 480
QY	481 GGGGATCTTCGGGATCTCTTCTCTCGTCTCCAGCTGTTTCAACATCTCGCTTCGCGG 540
DB	481 GGGGACCTTCGGGATCTGTCTTCTCTCGTCTCCAGCTGTTTCAACATCTCGCTTCGCGG 540
QY	541 CATGAGACGGTCAGGACTGCAATTTGCTCAATCTATCCGGGCAATAACGGGTACCGT 600
DB	541 CATGAGACGGTCAGGACTGCAATTTGCTCAATCTATCCGGGCAATAACGGGTACCGT 600
QY	601 ATGCTTGGGATATGATGATCACTGGT 628
DB	601 ATGCTTGGGATATGATGATCACTGGT 628
RESULT 6	
AA48914	
ID	AA48914 standard; DNA; 795 BP.
XX	AA48914;
XX	
DT	24-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus clone HCC110A E1 protein coding sequence.
XX	
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW	vaccine; immunostimulant; protein; vaccine; ds.
XX	
OS	Hepatitis C virus.
XX	
PN	WO20025548-A2.
XX	

18-JUL-2002.

11-JAN-2002; 2002WO-EP00219.

11-JAN-2001; 2001US-260699P.

30-AUG-2001; 2001US-315768P.

(INNO-) INNOGENETICS NV.

Maertens G, Bosman F, Buyse M;

WPI; 2002-599657/64.

P-PSDB; AAO18661.

New therapeutic vaccine compositions comprising at least one purified recombinant hepatitis C virus (HCV) single or specific oligomeric recombinant envelope protein E1 or E2, useful for immunizing humans from HCV infection -

Example 2; Page 161-162; 243pp; English.

The present invention relates to new therapeutic vaccine compositions for inducing hepatitis C virus (HCV)-specific antibodies, comprising a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope proteins selected from an E1 and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are useful for inducing HCV-specific antibodies or for immunising humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, or raising antibodies, in the preparation of medicament, and for in vitro monitoring of HCV disease or prognosing the response to treatment of patients suffering from HCV infection. The present sequence is a coding sequence described in the exemplification of the invention.

Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match	95.0%	Score 604	DB 24	Length 795
Best Local Similarity	97.6%	Pred. No. 3.3e-153		
Matches 613	Conservative 0	Mismatches 15	Indels 0	Gaps 0
Qy	1	ATGCTGGGTAAAGCCATCGATACCCCTTACGTGGGGCTTCGGGACCTCGTGGGGTACATT	60	
Db	1	ATGTTGGTAAAGGTACATCGATACCCCTTACATGGGGCTTCGGGACCTCGTGGGGTACATT	60	
Qy	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCAGAGGCCCTTGGGGCATGGGCTCCGGGTT	120	
Db	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTTGGGGCATGGGCTCCGGGTT	120	
Qy	121	CTGGAAGACGGCGTGAACTATGCAACAGGGAAATTGGCTGGTGTCTCTTTCTCTATCTTC	180	
Db	121	CTGGAGACGGCGTGAACTATGCAACAGGGAAATTTGCCCGGTTGCTCTTTCTCTATCTTC	180	
Qy	181	CTCTTGGCTTTACTGTCTGTCTTAAACATTTCCAGCTTCGCGTTACGAGGTGCGCAACGTG	240	
Db	181	CTCTTGGCTTTGTGTCCTGTCTGACCGTTCCAGCTTCGCGTTATGAAGTGCAGAACGTG	240	
Qy	241	TCGGGATGTACCATGTACAGAACGACTGCTCCAACCTCAAGCATTTGTATAGGCGACGC	300	
Db	241	TCGGGATGTACCATGTACAGAACGACTGCTCCAACCTCAAGCATTTGTATAGGCGACGC	300	
Qy	301	GACATGATCATGCACACCCCGGGTGGTGCCTTCGCGTTGGGAGAACAACTTTTCCCGC	360	
Db	301	GACATGATCATGCACACCCCGGGTGGTGCCTTCGCGTTGGGAGAACAACTTTTCCCGC	360	
Qy	361	TGCTGGGTAGCGGTACACCCACAGCTCGGGCTTAGGAACGCCAGCATCCCCACTACAAACA	420	
Db	361	TGCTGGGTAGCGGTACACCCACAGCTCGAGCTAGGAACGCCAGCGTCCCACCAACGACA	420	
Qy	421	ATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATCTACGTG	480	
Db	421	ATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATCTACGTG	480	
Qy	481	GGGGATCTCTGCGGATCTGTCTTCTCGTCTCCAGCTGTTCCACCATCTCGCCTCGCGCG	540	

RESULT 6	
AAL48914	
ID	AAL48914 standard; DNA; 795 BP.
XX	
XX	
AC	AAL48914;
XX	
DT	24-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus clone HCC110A E1 protein coding sequence.
XX	
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW	virucide; immunostimulant; vaccine; ds.
XX	
OS	Hepatitis C virus.
XX	
PN	WO20025548-A2.
XX	


```
PT produce proteins suitable for direct use, in vaccines or diagnostic
PT assays of HCV
XX
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or
XX the genotype of HCV. The new purification method preserves the
XX conformation of the recombinantly expressed E1, E2 and E1/E2, and
XX eliminates contaminating proteins. Antigens isolated using this method
XX are more reactive with human sera than those isolated by known
XX techniques.
XX Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;
XX
XX Query Match 94.6%; Score 601.6; DB 17; Length 2086;
XX Best Local Similarity 97.8%; Pred. No. 2e-152;
XX Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
XX
QY 5 TGGGTAAGGCCATCGATACCTTACGTGGGCTTCGCCGACCTCGTGGGGTACATTCGCG 64
DB |||||||
DB 5 TGGGTAAGGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGCG 64
QY 65 TCGTCCGCGCCCCCTAGGCGGCGCTCCAGGSCCTCGCGCATGCGTCCGGGTTCTGG 124
DB |||||||
DB 65 TCGTCCGCGCCCCCTAGGCGGCGCTCCAGGSCCTCGCGCATGCGTCCGGGTTCTGG 124
QY 125 AAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTATCTTCTCTCT 184
DB |||||||
DB 125 AAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTATCTTCTCTCT 184
QY 185 TGGCTTTACTGCTGCTCTAACTTCCAGCTTCGCGTTCAGAGGTGCGCAAGTGTCCG 244
DB |||||||
DB 185 TGGCTTTGCTGCTGCTCTGACCGTTCCAGCTTCGCGTTCAGAGGTGCGCAAGTGTCCG 244
QY 245 GGATGTACCATGTCACGACGACTGCTCAACTCAAGCATTTGTTATGAGCGGACG 304
DB |||||||
DB 245 GGATGTACCATGTCACGACGACTGCTCAACTCAAGCATTTGTTATGAGCGGACG 304
QY 305 TGATCATGCACACCCCCGGGTGCTGCCCTCGCTTCGGGAGAACAACTCTTCCCGTGTCT 364
DB |||||||
DB 305 TGATCATGCACACCCCCGGGTGCTGCCCTCGCTTCGGGAGAACAACTCTTCCCGTGTCT 364
QY 365 GGGTAGCGTTCACCCCAAGCTCGCGGTAGGAAACGCGACATCCCCACTTACAAACATAC 424
DB |||||||
DB 365 GGGTAGCGTTCACCCCAAGCTCGCGGTAGGAAACGCGACATCCCCACTTACAAACATAC 424
QY 425 GACGCCACCTCGAATTTGCTTGGGGCGGCTCTTCTGTTCCGCTATGTAGTGGGG 484
DB |||||||
DB 425 GACGCCACCTCGAATTTGCTTGGGGCGGCTCTTCTGTTCCGCTATGTAGTGGGG 484
QY 485 ATCTTCGCGGATCTGCTTCTCTGCTCCAGCTGTTTCAACATCTCGCTCGCGGCGATG 544
DB |||||||
DB 485 ACCTTCGCGGATCTGCTTCTCTGCTCCAGCTGTTTCAACATCTCGCTTCGCGGCGATG 544
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTTCACCGTATGG 604
DB |||||||
DB 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTTCACCGTATGG 604
QY 605 CTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
DB |||||||
DB 605 CTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
```

RESULT 9

```
AAT12974
ID AAT12974 standard; DNA; 2433 BP.
XX
XX AAT12974;
AC AAT12974;
XX
XX 25-SEP-1996 (first entry)
DT
XX HCV E1 construct HCC166.
DE
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
XX Hepatitis C virus.
OS
XX WO9604385-A2.
PN
XX 15-FEB-1996.
PD
XX 31-JUL-1995; 95WO-BP03031.
PF
XX 29-JUL-1994; 94EP-0870132.
PR
XX (INNO-) INNOGENETICS NV.
PA
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
PI
XX WPI; 1996-129401/13.
XX
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of disulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV.
XX
XX Claim 23; Fig 21; 146pp; English.
XX
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or
XX the genotype of HCV. The new purification method preserves the
XX conformation of the recombinantly expressed E1, E2 and E1/E2, and
XX eliminates contaminating proteins. Antigens isolated using this method
XX are more reactive with human sera than those isolated by known
XX techniques.
XX
XX Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;
XX
XX Query Match 94.6%; Score 601.6; DB 17; Length 2433;
XX Best Local Similarity 97.8%; Pred. No. 2.1e-152;
XX Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
XX
QY 5 TGGGTAAGGCCATCGATACCTTACGTGGGCTTCGCCGACCTCGTGGGGTACATTCGCG 64
DB |||||||
DB 356 TGGGTAAGGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGCG 415
QY 65 TCGTCCGCGCCCCCTAGGCGGCGCTCCAGGSCCTCGCGCATGCGTCCGGGTTCTGG 124
DB |||||||
DB 416 TCGTCCGCGCCCCCTAGGCGGCGCTCCAGGSCCTCGCGCATGCGTCCGGGTTCTGG 475
QY 125 AAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTATCTTCTCTCTCT 184
```


Db 476 AGGACGGCGTAACTATGCAACAGGAAATTTGCCCGGTGCTCTTTCTCTATCTCTCT 535
Qy 185 TGGCTTTACTGCTCTTAAACATTCACAGCTTCCGGCTTACGAGGTGCGCAACGTGTCG 244
Db 536 TGGCTTTGCTGCTGCTGACCGTTCCAGCTTCCGGCTTATGAAGTGGCAACGTGTCG 595
Qy 245 GGATGTACCATGTACGAAACGACTGCTCAACTCAAGCATTTGTATGAGCGACGGACA 304
Db 596 GGATGTACCATGTACGAAACGACTGCTCAACTCAAGCATTTGTATGAGCGACGGACA 655
Qy 305 TGNATGTCACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
Db 656 TGATCATGTACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
Qy 365 GGGTAGCGCTCACCCCGACGCTCGCGGTAGGAACGCCAGCATCCCACTACAAATAC 424
Db 716 GGGTAGCGCTCACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCACTACAAATAC 775
Qy 425 GAGCCACGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 484
Db 776 GAGCCACGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835
Qy 485 ATCTCTCGGATGCTCTTCTGCTCTCCAGCTGTTCAACATCTCGCTCGCGCATG 544
Db 836 ACCTCTGGGATGCTCTTCTGCTCTCCAGCTGTTCAACATCTCGCTCGCGCATG 895
Qy 545 AGACGGTGCAGGATGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
Db 896 AGACGGTGCAGGATGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 955
Qy 605 CTTGGGATGATGATGAACCTGGT 628
Db 956 CTTGGGATGATGATGAACCTGGT 979

RESULT 10
AAL48940
ID AAL48940 standard; DNA; 2434 BP.
XX AAL48940;
XX
DT 24-OCT-2002 (first entry)
XX Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.
DE Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.
XX Hepatitis C virus.
XX WO20025548-A2.
XX 18-JUL-2002.
XX 11-JAN-2002; 2002WO-EP00219.
XX 11-JAN-2001; 2001US-260699P.
PR 30-AUG-2001; 2001US-315768P.
XX
XX (INNO-) INNOGENETICS NV.
XX Maertens G, Boeman F, Buyse M;
XX WPI; 2002-599657/64.
DR P-PSDB; AAO18679.
XX
XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection -
XX
XX Example 2; Page 212-215; 243pp; English.

XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 other;

Query Match 92.9%; Score 590.6; DB 24; Length 2434;
Best Local Similarity 97.6%; Pred. No. 2e-149;
Matches 610; Conservative 0; Mismatches 14; Indels 1; Gaps 1;
Qy 5 TGGGTAAAGCCATGATACCTTACGTGCGGCTTCGCCGACCTCGTGGGTACATTCGCG 64
Db 356 TGGGTAAAGGTATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCGCG 415
Qy 65 TCGTCGGCGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGCGTTCG 124
Db 416 TCGTCGGCGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGCGTTCG 475
Qy 125 AAGAGCGGTGAATATGATGACAGGGAATTTGCCCTGCTGCTCTTCTCTATCTCTCT 184
Db 476 AGGACGGCGTGAATATGATGACAGGGAATTTGCCCTGCTGCTCTTCTCTATCTCTCT 535
Qy 185 TGGCTTTACTGTCCTG-TCTAACCATTTCCAGCTTCGCTTACGAGGTGCGCAAGTGTCC 243
Db 536 TGGCTTTGCTGCTGCTTTCACCGTTCCAGCTTCGCTTACGAGGTGCGCAAGTGTCC 595
Qy 244 GGGATGTACCATGTACGAAACGCTGCTCAACTCAAGCATTTGTATGAGCGACGGAC 303
Db 596 GGGATGTACCATGTACGAAACGCTGCTCAACTCAAGCATTTGTATGAGCGACGGAC 655
Qy 304 ATGATCATGCACACCCCGGGTGCCTGCGTTCGGGAGAACAACTCTTCCGCTGC 363
Db 656 ATGATCATGCACACCCCGGGTGCCTGCGTTCGGGAGAACAACTCTTCCGCTGC 715
Qy 364 TGGGTAGCGCTCACCCCGAGCTCGGGCTAGGAACGCGACATCCCACTACAAACAATA 423
Db 716 TGGGTAGCGCTCACCCCGAGCTCGAGCTAGGAACGCGAGCTCCCACTACAAACAATA 775
Qy 424 CGACGCCAGCTGATTTGCTTGGGGCGGCTGCTTCTGTTCCGCTATGATGCTGGG 483
Db 776 CGACGCCAGCTGATTTGCTTGGGGCGGCTGCTTCTGTTCCGCTATGATGCTGGG 835
Qy 484 GATCTCTCGGATCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
Db 836 GACCTCTGCGGATCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
Qy 544 GAGACGGTGCAGACTGCAATTCCTCAATCTATCCCGGCGCACATAACGGGTACCGGTATG 603
Db 896 GAGACGGTGCAGACTGCAATTCCTCAATCTATCCCGGCGCACATAACGGGTACCGGTATG 955
Qy 604 GCTTGGGATGATGATGAACCTGGT 628
Db 956 GCTTGGGATGATGATGAACCTGGT 980

RESULT 11
AAV42305
ID AAV42305 standard; cDNA; 673 BP.
XX
XX AAV42305;
XX
DT 26-OCT-1998 (first entry)
XX

12-JUN-1990; 90JP-0153401.
08-NOV-1990; 90JP-0304405.
(NAKA/) NAKAMURA T.
WPI; 1993-199637/25.
P-PSDB; AAR38279.
Antigen related to non-A and non-B hepatitis virus - comprises
non-translation region comprising 340 - 341 mols. of nucleotides,
PT non-translation region comprising 1885 - 2551 mols. of
PT nucleotides including region 1,149 and, etc.
XX Claim 3; Page 19-20; 73pp; Japanese.
PS The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
XX which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
CC The polypeptide it encodes may be used in a system for detecting
CC NANB hepatitis. This method is highly specific and sensitive, and
CC can detect NANB hepatitis virus which could not be detected by
CC conventional methods.
XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;
SQ Query Match 87.5%; Score 556.8; DB 14; Length 2540;
Beat Local Similarity 93.3%; Pred. No. 2.7e-140;
Matches 582; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 5 TGGGTAAGCCATCGATACCCCTTACGTGGGCTTCGCGACCTCGTGGGGTACATTCGCG 64
Db |||||
QY 356 TGGGTAAGGTCATCATGATACCCCTTACGTGGGCTTCGCGACCTCATGGGTATATTCGCG 415
Db |||||
QY 65 TCGTCGGCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGGTTCGGTTCGG 124
Db |||||
QY 416 TCGTCGGCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGGTTCGGTTCGG 475
Db |||||
QY 125 AAGACGGGTGAATGCAACAGGAAATTCGCTGGTTCCTTCTCTATCTTCCTCT 184
Db |||||
QY 476 AAGACGGGTGAATGCAACAGGAAATTCGCTGGTTCCTTCTCTATCTTCCTCT 535
Db |||||
QY 185 TGGCTTTACTGCTCTGCTTAACCATTCAGCTTCGCTTACGAGGTGCGCAACGTGTCG 244
Db |||||
QY 536 TGGCTTTGCTGCTGCTGTTGACCATCCAGCTTCGCTTATGAAGTGGCAACGTGTCG 595
Db |||||
QY 245 GGATGTACCATGTACGAACAGCTGCTCAACTCAAGCAATTTGTATGAGGCGGAC 304
Db |||||
QY 596 GGATATACCATGTACGAACAGCTGCTCAACTCAAGCAATTTGTATGAGGCGGAC 655
Db |||||
QY 305 TGATCATGCACACCCCGGCTGCGCTTCGCTTCGGGAGAACAACTCTTCCCGTGTCT 364
Db |||||
QY 656 TGATCATGCATACCTCCCGGTGCGCTTCGCTTCGGGAGAACAACTCTTCCCGTGTCT 715
Db |||||
QY 365 GGGTAGCGCTCACCCCAACGCTCGCGCTAGGAACGCGCAGCATCCCACTACAAATAC 424
Db |||||
QY 716 GGGTAGCGCTCACCTCCACGCTCGCGGCGAGGAATGCCAGCGTCCCACTACGACAATAC 775
Db |||||
QY 425 GACGCCAGCTCGATTGCTGTTGGGGGCGCTGCTTCTGCTTCGCTATGTAGCTGGGG 484
Db |||||
QY 776 GACGCCAGCTCGATTGCTGTTGGGGGCGCTGCTTCTGCTTCGCTATGTAGCTGGGG 835
Db |||||
QY 485 ATCTCTGGGATCTGCTTCTGCTGCTCCAGCTGTTTCACTATCTCGCTTCGCTTCGCTG 544
Db |||||
QY 836 ATCTCTGGGATCTGTTTCTGCTTCCAGCTGTTTCACTATCTCGCTTCGCTTCGCTG 895
Db |||||
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGCGCACATAAGCGGTACCGTATGG 604
Db |||||
QY 896 AGACAGTGCAGGACTGCAATTTGCTCAATCTATCCGCGCACATTTATCAGGTCACGCAATGG 955
Db |||||
QY 605 CTGGGATATGATGAACCTGGT 628
Db |||||
QY 956 CTGGGATATGATGAACCTGGT 979
Db |||||

RESULT 13
AAQ43889 standard; cDNA to mRNA; 2540 BP.
AC AAQ43889;
XX

21-OCT-1993 (first entry)

NANB hepatitis virus polynucleotide N-2540-2.

Non-A, non-B; virus; polymerase chain reaction; detection;

sensitive; specific; HCV; NANBH; ss.

Non-A, non-B hepatitis virus.

Key Location/Qualifiers

FT CDS 342..2540

FT 5'UTR 1..341

FT /*tag= a

FT /*tag= b

FT /note= "from 5' terminal of NANBH virus RNA"

XX JF05091884-A.

XX 16-APR-1993.

XX 10-APR-1991; 91JP-0196175.

XX

PR 12-JUN-1990; 90JP-0153401.
PR 08-NOV-1990; 90JP-0304405.
XX (NAKA/) NAKAMURA T.
XX WPI; 1993-199637/25.
XX P-PSDB; AAR38279.
XX Antigen related to non-A and non-B hepatitis virus - comprises
XX non-translation region comprising 340 - 341 mols. of nucleotides,
XX PT non-translation region comprising 1885 - 2551 mols. of
XX PT nucleotides including region 1,149 and, etc.
XX Claim 3; Page 19-20; 73pp; Japanese.
XX The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
XX which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
XX The polypeptide it encodes may be used in a system for detecting
XX NANB hepatitis. This method is highly specific and sensitive, and
XX can detect NANB hepatitis virus which could not be detected by
XX conventional methods.
XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;
SQ Query Match 87.5%; Score 556.8; DB 14; Length 2540;
Beat Local Similarity 93.3%; Pred. No. 2.7e-140;
Matches 582; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 5 TGGGTAAGCCATCGATACCCCTTACGTGGGCTTCGCGACCTCGTGGGGTACATTCGCG 64
Db |||||
QY 65 TCGTCGGCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGGTTCGGTTCGG 124
Db |||||
QY 757 TCGTCGGCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGGTTCGGTTCGG 816
Db |||||
QY 125 AAGACGGGTGAATGCAACAGGAAATTCGCTGGTTCCTTCTCTATCTTCCTCT 184
Db |||||
QY 817 AAGACGGGTGAATGCAACAGGAAATTCGCTGGTTCCTTCTCTATCTTCCTCT 876
Db |||||
QY 185 TGGCTTTACTGCTCTGCTTAACCATTCAGCTTCGCTTACGAGGTGCGCAACGTGTCG 244
Db |||||
QY 877 TGGCTTTGCTGCTGCTTACCATCCAGCTTCGCTTATGAAGTGGCAACGTGTCG 936
Db |||||
QY 245 GGATGTACCATGTACGAACAGCTGCTCAACTCAAGCAATTTGTATGAGGCGGAC 304
Db |||||
QY 937 GGATATACCATGTACGAACAGCTGCTCAACTCAAGCAATTTGTATGAGGCGGAC 996
Db |||||
QY 305 TGNATCATGCACACCCCGGCTGCGCTTCGCTTCGGGAGAACAACTCTTCCCGTGTCT 364
Db |||||
QY 997 TGNATCATGCATACCTCCCGGTGCGCTTCGCTTCGGGAGAACAACTCTTCCCGTGTCT 1056
Db |||||
QY 365 GGGTAGCGCTCACCCCAACGCTCGCGGTAGGAACGCGCAGCATCCCACTACAAATAC 424
Db |||||
QY 1057 GGGTAGCGCTCACCTCCAGCTCGCGGCGAGGAATGCCAGCGTCCCACTACAAATAC 1116
Db |||||
QY 425 GACGCCAGCTCGATTGCTGTTGGGGGCGCTGCTTCTGCTTCGCTATGTAGTGGGG 484
Db |||||
QY 1117 GACGCCAGCTCGATTGCTGTTGGGGGCGCTGCTTCTGCTTCGCTATGTAGTGGGG 1176
Db |||||
QY 485 ATCTCTGGGATCTGCTTCTGCTTCGAGCTGTTTCACTATCTCGCTTCGCTTCGCTG 544
Db |||||
QY 1177 ATCTCTGGGATCTGTTTCTGCTTCGAGCTGTTTCACTATCTCGCTTCGCTTCG 1236
Db |||||
QY 545 AAGACGGGTGCAGGACTGCAATTTGCTCAATCTATCCGCGCACATAAGCGGTACCGTATGG 604
Db |||||
QY 1237 AAGACGGGTGCAGGACTGCAATTTGCTCAATCTATCCGCGCACATTTATCAGGTCACCGCATGG 1296
Db |||||
QY 605 CTGGGATATGATGAACCTGGT 628
Db |||||
QY 1297 CTGGGATATGATGAACCTGGT 1320
Db |||||


```
RESULT 14
AAQ63753
ID AAQ63753 standard; cDNA to mRNA; 2540 BP.
XX
AC AAQ63753;
XX
DT 30-JAN-1995 (first entry)
XX
DE NANBHV genomic fragment #2.
XX
XX Polymerase chain reaction; PCR; primer; amplify; detection; NANBHV;
KW non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.
XX
OS Synthetic.
XX
PN JP06125777-A.
XX
PD 10-MAY-1994.
XX
PF 20-JUN-1991; 91JP-0247120.
XX
PR 20-JUN-1991; 91JP-0247120.
XX
PA (NAKA/) NAKAMURA T.
XX
DR WPI; 1994-187937/23.
XX
XX Oligonucleotide primer pairs specific for non-A, non-B hepatitis
PT virus - used for high sensitivity detection of non-A non-B (NANB)
PT hepatitis virus
XX
PS Disclosure; Page 24-25; 25pp; Japanese.
XX
CC The sequences given in AAQ63752-53 represent fragments of the non-A,
CC non-B hepatitis virus (NANBH) genome. These fragments were amplified
CC using the primers given in AAQ63732-51. These primers were used in the
CC detection of NANBH. The primers are based on the 5'-terminal region and
CC the core protein coding region. The method allows highly sensitive
CC detection of NANBH.
XX
SQ Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 other;

Query Match 87.5%; Score 556.8; DB 15; Length 2540;
Best Local Similarity 93.3%; Pred. No. 2.7e-140;
Matches 582; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 5 TGGGTAAGGCCATCGATACCTTACCTGGGGCTTCGCCGACCTCGTGGGGTACATTCGCG 64
DB |||||||
DB 697 TGGGTAAGGTCATCGATACCTTACATGCGGCTTCGCCGATCTCATGGGGTATATTCGCG 756
QY 65 TCGTCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGG 124
DB |||||||
DB 757 TCGTCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGCACACGGGTGCCGGGTTCTGG 816
QY 125 AAGACGGCGTGAATATGCAACAGGGAATTTGGCTGGTCTTCTCTATCTCTCTCT 184
DB |||||||
DB 817 AGACGGCGTGAATATGCAACAGGGAATTTGGCTGGTCTTCTCTATCTCTCTCTCT 876
QY 185 TGGCTTTACTGCTCTCTAAACCATTTCCAGCTTCGCTTACGAGGTCGCAAGCTGTCGG 244
DB |||||||
DB 877 TGGCTTTGCTGCTCTGTTTACCATCCAGCTTCCGCTTATGAAGTGGCAAGGTGCG 936
QY 245 GGATGTACCATGTACGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCGCGGACA 304
DB |||||||
DB 937 GGATATACCATGTACGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCGCGGACA 996
QY 305 TGATCATGACACCCCGGGTGGTCCCTCGCTTGGGAGAACACTCTTCCCGTGTCT 364
DB |||||||
DB 997 TGATCATGATATCTCCCGGTGGTGGCTTCCGCTTCCGGAGGACACAGCTCCCGTGTCT 1056
QY 365 GGGTAGCGCTCACCCCGCTCGGGTAGGACGCGGATCCCGCTACACAAATAC 424
DB |||||||
DB 1057 GGGTAGCGCTCACTCCCGCTCGGGCAGGAATCCAGCGTCCCGCTACACAAATAC 1116
```

```
QY 425 GACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATAGTGGGG 484
DB |||||||
DB 1117 GACGCCACGTCGACTTGTCTGTTGGGGCGGCTGCTTTCTGCTCCGCTATGACGTGGGG 1176
QY 485 ATCTCTGCGGATCTGCTTCTCTGCTCTCCAGCTGTTCCACATCTCGCCTCGCGGCATG 544
DB |||||||
DB 1177 ATCTCTGCGGATCTGTTTCTCTGCTCTCCAGCTGTTCCACCTTCTCGCCTCGCGGCATG 1236
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGG 604
DB |||||||
DB 1237 AGACAGTGCAGGACTGCAACTGCTCAATCTATCCCGGCCATTTATCAGGTACCGCATGG 1296
QY 605 CTTGGGATATGATGATGAAGTGT 628
DB |||||||
DB 1297 CTTGGGATATGATGATGAAGTGT 1320

RESULT 15
ABK91411
ID ABK91411 standard; DNA; 9605 BP.
XX
AC ABK91411;
XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..9374
FT /tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
XX
XX WO200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-BP00526.
XX
XX 23-JAN-2001; 2001US-263479P.
XX
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI; 2002-599793/64.
XX
XX P-PSDB; ABG32451.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression
XX
XX Claim 9; Page 36-39; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
```


Mon Dec 22 13:28:34 2003

HCV (hepatitis C virus) replicon enhanced cell or which containing a functional HCV replicon; (5) an HCV replicon enhanced cells made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicons and HCV replicon enhanced cells are useful in studying HCV replication and expression, and HCV and host cell interactions, producing HCV RNA and proteins, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated diseases such as liver failure, cirrhosis and hepatocellular carcinoma. The present sequence is the HCV replicon Con 1, used as a basis for the presentive mutations of the invention.

XX
C: 2733 G: 2079 T: 0 other;
C: 2883 C: 1910 A: 2005 P: 0

Query Match	87.5%	Score	556.8	DB	24	Length	9605
Best Local Similarity	93.3%	Pred.	No. 4e-140	Indels	0	Gaps	0
Matches	582	Conservative	0	Mismatches	42	Indels	0
QY	5	TGGTGAAGGCATCGATACCCCTTACGCTGCGGGCTTCGCCAACCTCGTGGGGTACATTCGCG	64				
Db	697	TGGGTAGGTCAATGATACCTCAGCTGCGGCTTCGCGATCTCATGGGGTACATTCGCG	756				
QY	65	TCGTCGCGGCCCTTACGGGGGCGCTCCAGGGCCCTGGCGATGCGCTCCGGTTCGCG	124				
Db	757	TCGTCGCGGCCCTTACGGGGGCGCTCCAGGGCCCTGGCGATGCGCTCCGGTTCGCG	816				
QY	125	AAGACGGGTCAACTATGCAACAGGAATTTGCTGGTTGCTCTCTTTCTATCTTCCTCT	184				
Db	817	AGGACGGCGTGAACATAAGCAACAGGAATCTGCCCGGTTGCTCTTTCTATCTTCCTTT	876				
QY	185	TGGCTTTTACTGTCCTGTCTAAACCAATCCAGCTTCGCTTACGAGGTGCGCAACGTCTCG	244				
Db	877	TGGCTTTTCTGCTCTGTTTGAACATCCAGCTTCGCTTATGAAGTTCGCAACGTATCCG	936				
QY	245	GGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCATTTGTATAGCGCAGCGACA	304				
Db	937	GAGTGTACCATGTACAGAACGACTGCTCCAAACGACGATTTGTATAGCGCAGCGACA	996				
QY	305	TGATCATGCACACCCCGGGTGCCTGCGCTTCGGTTCCGGAGAACAACTCTTCGCGTGCT	364				
Db	997	TGATCATGCATACCCCGGGTGCCTGCGCTTCGGTTCCGGAGAACAACTCTTCGCGTGCT	1056				
QY	365	GGGTAGCGCTACCCCGCAGCTCGCGGTAGGAACGCCAGCATCCCCACATACAAATAC	424				
Db	1057	GGGTAGCGCTACTCCACGCTCGCGGCCAGGAACGCTAGCGTCCCACTACGACGATAC	1116				
QY	425	GACGCCACCTCGAATTGCTTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGTACGTGGGG	484				
Db	1117	GACGCCATGTGATTTGCTTCGTTGGGGGGCTGCTCTCTGCTCCGCTATGTACGTGGAG	1176				
QY	485	ATCTTCGGGATCTGCTTCCTCGTCTCCAGCTGTTACCAATCTCGCTCCCGGGCATG	544				
Db	1177	ATCTTCGGGATCTGTTTTCTCGTTCGCGCAGCTGTTACCTTCTTCGCTCCCGGCACG	1236				
QY	545	AGACGTTGACGAGTCAATTTGCTCAATCTATTCGCGGCCACATAACGGGTACCGTATGG	604				
Db	1237	AGACGTTACGAGTCAATTTGCTCAATATATCCCGGCCACGTGACAGTACCGCTATGG	1296				
QY	605	CTTGGGATATGATGAATGTT	628				
Db	1297	CTTGGGATATGATGAATGTT	1320				

Search completed: December 19, 2003, 18:51:08
Job time : 177.828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1636.15 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-13
Perfect score: 636
Sequence: 1 ATGCTGGTAAAGCCATCGA.....TGATGAACCTGGTACTAATAG 636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estnu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pin.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_pbg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	75.4	11.9	488	9 AV755731	AV755731
C 2	61.	9.6	492	9 AV758366	AV758366
C 3	42	6.6	1201	13 BX356664	BX356664
C 4	41.6	6.5	534	14 CD040840	CD040840

C 5	41.4	6.5	1270	12	BG968359	BG968359
C 6	41	6.4	1201	13	BX381961	BX381961
C 7	40.4	6.4	1097	9	AL580130	AL580130
C 8	40.2	6.3	1201	9	AL513886	AL513886
C 9	39.6	6.2	925	29	CNS0091P	AL053013 PROSOPHIL
C 10	39	6.1	359	12	BJ252669	BJ252669
C 11	39	6.1	375	12	BJ246716	BJ246716
C 12	39	6.1	621	14	CAB16001	CAB16001
C 13	39	6.1	856	29	BZ578381	BZ578381
C 14	39	6.1	872	29	BZ555011	BZ555011
C 15	39	6.1	1201	13	BX460099	BX460099
C 16	38.8	6.1	533	29	CC010084	CC010084
C 17	38.8	6.1	659	29	CC405164	CC405164
C 18	38.8	6.1	826	29	BZ736582	BZ736582
C 19	38.8	6.1	895	29	CC359028	CC359028
C 20	38.8	6.1	925	29	CC359026	CC359026
C 21	38.8	6.1	940	29	CC010085	CC010085
C 22	38.8	6.1	951	29	CC405167	CC405167
C 23	38.6	6.1	399	9	AV638521	AV638521
C 24	38.6	6.1	434	9	AV637507	AV637507
C 25	38.6	6.1	440	9	AV637983	AV637983
C 26	38.6	6.1	450	9	AV637259	AV637259
C 27	38.6	6.1	451	9	AV637328	AV637328
C 28	38.6	6.1	451	9	AV637643	AV637643
C 29	38.6	6.1	453	9	AV634724	AV634724
C 30	38.6	6.1	454	9	AV637050	AV637050
C 31	38.6	6.1	456	9	AV635382	AV635382
C 32	38.6	6.1	473	9	AV632765	AV632765
C 33	38.6	6.1	481	9	AV635503	AV635503
C 34	38.6	6.1	485	9	AV632811	AV632811
C 35	38.6	6.1	506	9	AV392445	AV392445
C 36	38.6	6.1	508	9	AV634095	AV634095
C 37	38.6	6.1	526	9	AV641895	AV641895
C 38	38.6	6.1	533	9	AV638125	AV638125
C 39	38.6	6.1	537	9	AV632335	AV632335
C 40	38.6	6.1	588	9	AV387329	AV387329
C 41	38.6	6.1	945	29	CNS05KMQ	AL341675 Tetraodon
C 42	38.4	6.0	525	10	BE337089	BE337089
C 43	38.4	6.0	671	12	BI723733	BI723733
C 44	38.2	6.0	431	9	AV639153	AV639153
C 45	38.2	6.0	501	9	AV638474	AV638474

ALIGNMENTS

RESULT 1
AV755731/c 488 bp mRNA linear EST 19-OCT-2000
LOCUS AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones

Unpublished

CONTACT: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919 (ex. 45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

Location/Qualifiers

FEATURES


```

1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="BMFAKB03"
/clone="BMFAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site_1: sfIIA; Site_2: sfIIIB"
116 a 134 c 137 g 97 t 4 others

BASE COUNT
ORIGIN

Query Match 11.9%; Score 75.4; DB 9; Length 488;
Best Local Similarity 70.3%; Pred. No. 3.6e-09;
Matches 130; Conservative 0; Mismatches 51; Indels 4; Gaps 2;

445 GTTGGGGCGGCTTCCTTCGTTCCGCTATGATGACGTGGGGATCTCTGGCGATCTGCTTC 504
|||
472 GTGGGTGTCACTCGCGCTTGCTCAGCTCTCTACGTTGGGACCTCTGCGACGGAGTGATG 413
|||
505 CTCGTCCTCCAGCTGTTCACATCTCGCTCGCCGGCATGACCGTGCAGACCTGCAAT 564
|||
412 CTTGCAGTTTCAGCTG--ATCATCTCGCCTCAGCACCATGAGTTGTGCATGAATGCAC 356
|||
565 TGCCTCAATCTATCCCGGCCACATAACCGGTTCACCGTATG- GCTTGGGATATGATGATA 623
|||
355 TGTCTCATCTATCTCTGGGCCCATCACTGGACACCGGTATGAGCATGGGACATGATGATA 296
|||

624 CTGGT 628
|||||
295 CTGGT 291
|||||
QY Db

```

RESULT 2
AV758366/c
LOCUS
AV758366 BM Homo sapiens cDNA clone BMFAXA03 5', mRNA linear EST 19-OCT-2000
DEFINITION
AV758366
ACCESSION
AV758366.1 GI:10916214
VERSION
EST.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
TITLE
Homo sapiens cDNA BM clones
COMMENT
Unpublished
Contact: Zeqiang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzq@chgc.sh.cn
This clone is available at CHGC in Shanghai.

```

FEATURES
source
1. .492
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAKA03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
124 a 128 c 125 g 112 t 3 others
BASE COUNT
ORIGIN

```

Query Match	9.6%	Score 61;	DB 9;	Length 492;
Best Local Similarity	68.0%	Pred. No. 2.1e-05;	Indels 4;	Gaps 2;
Matches 115;	Conservative 0;	Mismatches 50;		
461	TCCTGTCGGTATGATGACGTGGGGATCTCTGCGGATCTCTTCTCTCGTCTCCAGCTGT	520		
457	TGTGATCAGCTCACTACGTGTTGGACCTCTGCGTTGGGGTATCGTTGCAGCCCACTG-	399		
521	TCACCATCTCGCCTCGCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCG	580		
398	--ATACTCTCAGCAGCAACATTGGTTGTGCAGAAATGCAACTGCTCATTTCTCCTG	341		
581	GCCACATAACGGGTCT-ACCGTATGCGCTGGGATATGATGATGAACCTGGT	628		
340	GCTGCATCACTGGACTACAGTATGGCATAGGCTATGATGATGAACCTGGT	292		
RESULT 3				
LOCUS	BX356664			
DEFINITION	BX356664 Homo sapiens LINEA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI015YB03 3-PRIME, mRNA sequence.			
ACCESSION	BX356664			
VERSION	BX356664.1	GI:30378083		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished			
COMMENT	Contact: Genoscope			
	Genoscope - Centre National de Sequencage			
	BP 191 91006 EVRY cedex - France			
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
	Library was constructed by Life Technologies, a division of			
	Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :			
	http://fulllength.invitrogen.com/Invitrogen Corporation 1600			
	http://fulllength.invitrogen.com/Invitrogen Corporation 1600			

[illegible]

Db 1125 AGTGGCCCTTGGTGGTCTCTCTCTGGCGGTGACATCCTGGCGGTCTGGCCCTC 1066

QY 141 TGAACAGAGGAATGGCTGGTGTCTTTCTCTATCTCTCTGGTCTTACCTGCTG 200

Db 1065 TGCCTGACATCTGTCTCTGGTGATACATCTCCGCTGTCCCGGAGTGGTCTTGTCCCG 1006

QY 201 TCTAACCATTCAGCTTCCGCTTACGAGTGGCAACGCTGCGGGATGTACCATGTCA 260

Db 1005 TCCATCTCCCGGATCTCTCTGGTGGTCCGATCTCCGCTGTCCCGGAGTGGTCTTGTCCCG 946

QY 261 GAACGACTCTCCAACTCAAGCATTTGTGATGAGGAGGAGCATGATCATGCACACCC 320

Db 945 CATCCCTCTGTATCTTGTATGGCTCTCTGGCTGTCCCGGAGTGGTCTTGTCCCG 886

QY 321 CGGGTGGCTGCTCTGGTGGGAGCAACTCTCCGCTGTCTGGTGGGCTCACCCC 380

Db 885 CGGGTGGCTCTCAAGGCTCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 826

QY 381 CACGCTCGGGCTAGGAAGCCAGCATCCCC 411

Db 825 TTGCTTGGCTCTTGGGTACCTCTTCCGC 795

RESULT 6

BX381961/c 1201 bp mRNA linear EST 08-MAY-2003

LOCUS BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

DEFINITION clone CSODI072YF05 3-PRIME, mRNA sequence.

ACCESSION BX381961

VERSION BX381961.1 GI:30453007

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1201)

TITLE Li, W.B., Gruber, C., Jesses, J. and Polayes, D.

JOURNAL Full-length cDNA libraries and normalization

COMMENT Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSODI072CC03NP1.

Location/Qualifiers

1. .1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI072YF05"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 95 a 191 c 115 g 55 t 745 others

ORIGIN

Query Match 6.4%; Score 41; DB 13; Length 1201;

Best Local Similarity 3.3%; Pred.No. 4.8;

Matches 19; Conservative 189; Mismatches 369; Indels 0; Gaps 0;

QY 30 GTGCGCTTCCGCGACTCGTGGGGTACATTCGGCTGCTGCGCGCCCTAGGGGGCGC 89

Db 970 GRWVKBMGKMKMKKBMKBNKKTMMWMBKMKMKMKMKMKMKMKMKMKMKMKMKMKMK 911

QY 90 TGGCAGGGCCCTGGCGCATGGCTCGGGTCTCGGAGACGGCGTGAACATGCAACAGG 149

Db 910 MMKKGKMK 851

QY 150 GAATTTGGCTGGTGGTCTTTCTCTATCTCTCTGGCTTCTCTCTCTCTCTCTCTCTCT 209

Db 850 KMMVKGKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMK 791

QY 210 TCCAGCTTCCGCTTACGAGTGGCAACGCTGTCGGGATGTACCATGTGTCAGAACGACTG 269

Db 790 MSKKKNN 731

QY 270 CTCCAACTCAAGCAATTTGTATGAGGACGGGACATGATCATGCACACCCCGGGTGGCT 329

Db 730 KMMVNN 671

QY 330 GCCCTGGTTCGGGAGCAACACTTCTCCCGCTGCTGGGTAGCGCTCACCCCGACGCTCGC 389

Db 670 NNN 611

QY 390 GGCTAGGAACGCCAGCATCCCACTACAACTACGACGCCAGCTGATGCTGCTCTGGTGG 449

Db 610 NNN 551

QY 450 GCGCGCTGCTTCTGTTCCGCTATGTACGTTGGGGATCTCTCGGATCTGCTCTCTCTCGT 509

Db 550 MNKKKNN 491

QY 510 CTCCAGCTGTTTACCACTCTCGCTCGCGGCGATGAGACGGTGCAGGACTGCAATGCTC 569

Db 490 NNN 431

QY 570 AATCTATCCCGGCCACATAACGGGTACCGGTATGGCT 606

Db 430 MNKKKNN 394

RESULT 7

AL580130/c 1097 bp mRNA linear EST 01-JUN-2003

LOCUS AL580130 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED

DEFINITION Homo sapiens cDNA clone CSODJ001YK17 3-PRIME, mRNA sequence.

ACCESSION AL580130

VERSION AL580130.2 GI:31318409

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1097)

TITLE Li, W.B., Gruber, C., Jesses, J. and Polayes, D.

JOURNAL Full-length cDNA libraries and normalization

COMMENT Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 874.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CSODJ001AF09NP1&cluster=874.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSODJ001AF09NP1.

Location/Qualifiers

1. .1097

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODJ001YK17"

/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

/cell_line="JURKAT"

/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

FEATURES

source


```
source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match 6.2%; Score 39.6; DB 29; Length 925;
Best Local Similarity 12.1%; Pred. No. 10;
Matches 45; Conservative 164; Mismatches 163; Indels 0; Gaps 0;

Qy 2 TGTGTTGTTAGCCATCATACCTTACGTGGGCTTCGCCACCTCGTGGGGTACATTC 61
Db 553 TTSSGGYGGKSSGGBSCSCSSSCSSSCSCBCCCCSCSSYCCSSBSKSSKSS 612

Qy 62 CGCTCGTCGGCGCCCTAGGGCGCTGCAGAGCCCTGGCGCATGCGTCCGGTTC 121
Db 613 TSBSGCCSKSVCGTSCSSSSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 672

Qy 122 TGAAGAGCGCGTGAATGCAACAGGGAATTGCTGTTGCTTTCTTCTCTATCTCC 181
Db 673 KSTASGSGWSAGSGSGSTGTSSTSSSSSTSSSTSSSSKSTSSBSSBSSSGSS 732

Qy 182 TCTTGGCTTTACTGCTGCTGTAACATTCACGTTCCGCTTACGAGGTGCGCAACGTG 241
Db 733 SSSTSSBBSCTSSSSSSSSSTSCCTCCCSYSSSTSSSTSSSTSSSTSSSTSSSV 792

Qy 242 CCGGAGTACCATGTGCAGAACAGTCTCCAACTCAAGCATGTTGTTAGGCGCGG 301
Db 793 GTSSSSDSTSTCCGCCYMCCTCCSTYBMCYTSTSCGSSSSSGKGGVTKCGCGGSS 852

Qy 302 ACATGATCATGCACACCCCGGTGCTGCGTTCGGGAGAACAACTCTTCCCGCT 361
Db 853 TNGMBGTSSACSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSV 912

Qy 362 GCTGGTAGCGC 373
Db 913 SSGSGSGSVS 924

RESULT 10
BU252669/c 359 bp mRNA linear EST 05-APR-2002
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 3', mRNA sequence.
ACCESSION
BU252669.1 GI:20061830
VERSION
EST.
KEYWORDS
Triticum aestivum (bread wheat)
SOURCE
Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 359)
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Peekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
/clone 107 c 110 g 77 t

BASE COUNT 81 a 107 c 110 g 77 t
ORIGIN

Query Match 6.1%; Score 39; DB 12; Length 375;
Best Local Similarity 58.0%; Pred. No. 11;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 24 CCTTACGTGGGCTTCGCCACCTCGTGGGGTACATTCGCTCGTGGCGCCCTTAGG 83
Db 36 CTTCAAGTGCACAGCGCGCTCTGGAAGCGCTCTGGAAGCGCTCAGGGCGGTGCGCGTCGG 95

Qy 84 GGGCGCTGCCAGGCGCTGCGCATGCGTCCGGTTCGGAAGACGGCGTGAACATG 142
Db 96 GGACCGCGCAGCGCCCTGGCGCAGGACGTGCGAGTCTGCGCGTGCACGTGCCCAAG 154

RESULT 11
BU246716 375 bp mRNA linear EST 05-APR-2002
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 5', mRNA sequence.
ACCESSION
BU246716
VERSION
EST.
KEYWORDS
Triticum aestivum (bread wheat)
SOURCE
Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 375)
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..375
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Peekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
/clone 107 c 110 g 77 t

BASE COUNT 81 a 107 c 110 g 77 t
ORIGIN

Query Match 6.1%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 11;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 24 CCTTACGTGGGCTTCGCCACCTCGTGGGGTACATTCGCTCGTGGCGCCCTTAGG 83
Db 297 CTTCAAGTGCACAGCGCGCTCTGGAAGCGCTCTGGAAGCGCTCAGGGCGGTGCGCGTCGG 238

Qy 84 GGGCGCTGCCAGGCGCTGCGCATGCGTCCGGTTCGGAAGACGGCGTGAACATG 142
Db 237 GGACCGCGCAGCGCCCTGGCGCAGGACGTGCGAGTCTGCGCGTGCACGTGCCCAAG 179

RESULT 12
CA816001 621 bp mRNA linear EST 09-DEC-2002
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 5', mRNA sequence.
ACCESSION
CA816001
VERSION
EST.
KEYWORDS
Triticum aestivum (bread wheat)
SOURCE
Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 621)
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..621
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Peekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
/clone 115 c 107 g 67 t

BASE COUNT 70 a 115 c 107 g 67 t
ORIGIN
```


DEFINITION CAL2EI1011VF_E04 Cabernet Sauvignon Leaf - CAL2EI Vitis vinifera
 CDNA clone CAL2EI1011VF_E04 5', mRNA sequence.
 ACCESSION CAB16001
 VERSION CAB16001.1 GI:26264938
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; Vitaceae; Vitis.
 REFERENCE 1 (bases 1 to 621)
 AUTHORS Goes da Silva,F., Lim,H., Iandolino,A., Baek,J., Jones,K., Walker
 ,M.A. and Cook,D.R.
 TITLE Transcriptional responses of Vitis vinifera to infection by the
 bacterial pathogen Xylella fastidiosa
 JOURNAL Unpublished
 COMMENT Contact: Doug Cook
 CAES Genome Facility
 UC Davis Department of Plant Pathology
 1 Shields Ave., Davis, CA 95616, USA
 Tel: 530 754 8561
 Fax: 530 754 6617
 Email: drcook@ucdavis.edu
 Seq primer: GTTATCAGTCGACGGTACC.

FEATURES

source
 1..621
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAL2EI1011VF_E04"
 /sex="hermaphrodite"
 /dev_stage="Mid-season leaf material"
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Leaf - CAL2EI"
 /note="Organ: Leaf. Vector: pDNR. Site 1: Sfil; Site 2:
 Sfil; CAL2EI is a cDNA library of Cabernet Sauvignon
 leaves. The leaves were collected on July 25, 2001, in
 Napa Valley, California, and represent leaves in
 mid-season development. These leaves were verified to be
 infected with the bacterial pathogen, Xylella fastidiosa,
 based on a diagnostic assay using PCR and Xylella-specific
 primer pairs. The plants were asymptomatic at the time of
 collection, but later developed symptoms. cDNAs were made
 by oligo-dT priming and directionally cloned. 5' and 3',
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGACAGTGGCATTACGCCGGG-3' and
 5'-ATTCTAGCGCGAGCGGCCGACATG-dT(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT

146 a 152 c 143 g 180 t

Query Match 6.1%; Score 39; DB 14; Length 621;

Best Local Similarity 50.3%; Pred. No. 13;
 Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 264 CGACTGCTCCAACTCAAGCATTTGTATGAGGAGCGGACATGATCATGCACACCCCGG 323
 |||||
 DB 397 CGACTTGTCTTCTATGACGGGTAGCGGACGAGGACATCCGAGTCTCCGAG 456
 |||||
 QY 324 GTGCGTGCCTCGTTTCGGGAGAACATCTTTCCGCTGCTGGTAGCGCTCACCCCCAC 383
 |||||
 DB 457 GTGCTCGAGATGCTTGTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 516
 |||||
 QY 384 GTCGCGGTAGAACCGCAGATCCCACTACAACATACGACGCCAGCTCGATTGCT 443
 |||||
 DB 517 ACACCTTCGGTGTTCGGCAGCGTAGCCGCTGACACCATCGAAGCATAGCTGGTTCGG 576
 |||||
 QY 444 CGTTGGGCGG 454
 |||||
 DB 577 AGTAGTGGGG 587

RESULT 13

BZ578381

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..856

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="MSH"

/db_xref="taxon:287"

/clone="msh2 5817"

/clone_lib="msh"

/note="Environmental isolate. Whole genomic shotgun
 library."

BASE COUNT 147 a 274 c 279 g 154 t

ORIGIN

Query Match 6.1%; Score 39; DB 29; Length 856;

Best Local Similarity 58.0%; Pred. No. 14;

Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 287 TGTATGAGCAGCGGACATGATCATGCACACCCCGGTCGTCGCTTCGGGAGA 346

|||||

DB 253 TTTTGGCGCGCGGACAGCCATCGCAGCGCGGTTCTCTCGACGAAGC 312

|||||

QY 347 ACACTCTTCCCGTGTGGGTAGCGCTACCCCGCTCGCGGTAGGAAGCCAGC 405

|||||

DB 313 CCCACTCGGTGCGCAGCAGGTCGCGCTTCGCTCGCGGTAGAGTCCAGC 371

|||||

RESULT 14

BZ555011

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 872)

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="MSH"

/db_xref="taxon:287"

/clone="msh2 5817"

/clone_lib="msh"

/note="Environmental isolate. Whole genomic shotgun
 library."

BASE COUNT 147 a 274 c 279 g 154 t

ORIGIN

Query Match 6.1%; Score 39; DB 29; Length 856;

Best Local Similarity 58.0%; Pred. No. 14;

Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 287 TGTATGAGCAGCGGACATGATCATGCACACCCCGGTCGTCGCTTCGGGAGA 346

|||||

DB 253 TTTTGGCGCGCGGACAGCCATCGCAGCGCGGTTCTCTCGACGAAGC 312

|||||

QY 347 ACACTCTTCCCGTGTGGGTAGCGCTACCCCGCTCGCGGTAGGAAGCCAGC 405

|||||

DB 313 CCCACTCGGTGCGCAGCAGGTCGCGCTTCGCTCGCGGTAGAGTCCAGC 371

|||||


```

1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDF010XP12"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dr) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
188 a 292 c 134 g 110 t 417 others

```

Search completed: December 20, 2003, 06:54:45
Job time : 1638.15 secs

Search completed: December 20, 2003, 06:54:45
Job time : 1638.15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 45.8681 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-13

Perfect score: 636

Sequence: 1 ATGCTGGTAAAGCCATCGA.....TGATGACTGGTACTAATAG 636

Scoring table: IDENTITY_NUC

Gapop 10.0 , Capext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCUTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	636	100.0	636	3	US-08-612-973-13
2	636	100.0	636	3	US-08-927-597-13
3	612.8	96.4	633	3	US-08-612-973-7
4	612.8	96.4	633	3	US-08-927-597-7
5	604	95.0	795	3	US-08-612-973-5
6	604	95.0	795	3	US-08-927-597-5
7	601.6	94.6	2082	3	US-08-612-973-47
8	601.6	94.6	2082	3	US-08-927-597-47
9	601.6	94.6	2433	3	US-08-612-973-49
10	601.6	94.6	2433	3	US-08-927-597-49
11	553.6	87.0	1539	2	US-08-470-426B-17
12	553.6	87.0	1863	2	US-08-470-426B-14
13	550.4	86.5	932	1	US-08-081-072-15
14	550.4	86.5	932	1	US-08-449-093A-15
15	548.8	86.3	1037	1	US-08-462-195-1
16	548.8	86.3	1037	3	US-08-636-883-1
17	548.8	86.3	1037	3	US-09-127-829-1
18	548.8	86.3	2116	3	US-08-191-160-21
19	545.6	85.8	9595	3	US-09-014-416-4
20	545.6	85.8	9599	3	US-09-014-416-6
21	534.4	84.0	1167	1	US-08-324-977-9
22	534.4	84.0	1167	2	US-08-384-616-9
23	534.4	84.0	1167	2	US-08-904-686A-9
24	534.4	84.0	1167	3	US-09-315-850-9
25	534.4	84.0	1499	1	US-08-324-977-3
26	534.4	84.0	1499	2	US-08-384-616-3
27	534.4	84.0	1499	2	US-08-904-686A-3

Sequence 3, Appli
Sequence 11, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 27, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-13
; Sequence 13, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..633
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..630

Mon Dec 22 13:28:35 2003

```
US-08-612-973-13
Query Match      100.0%; Score 636; DB 3; Length 636;
Best Local Similarity 100.0%; Pred. No. 3.3e-164;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/612,973
; APPLICATION NUMBER:
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630
;
; US-08-927-597-13
;
; Query Match      100.0%; Score 636; DB 3; Length 636;
; Best Local Similarity 100.0%; Pred. No. 3.3e-164;
; Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 ATGCTGGGTAAGCCATCGATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATT 60
; DB 1 ATGCTGGGTAAGCCATCGATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATT 60
;
; QY 61 CCGCTCGTGGGCGCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120
; DB 61 CCGCTCGTGGGCGCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120
;
; QY 121 CTGGAAGACGGGTGAACCTATGCAACAGGGAATTTGCCCTGCTTCTTCTATCTTC 180
; DB 121 CTGGAAGACGGGTGAACCTATGCAACAGGGAATTTGCCCTGCTTCTTCTATCTTC 180
;
; QY 181 CTCTTGGCTTTACTGCTCTTAACTTCCAGCTTCCGCTTACGAGTGGCGCAACGTG 240
; DB 181 CTCTTGGCTTTACTGCTCTTAACTTCCAGCTTCCGCTTACGAGTGGCGCAACGTG 240
;
; QY 241 TCCGGGATGTACCATGTCAAGAACGATGCTCCAATCAAGCAATTTGTATGAGGACGG 300
; DB 241 TCCGGGATGTACCATGTCAAGAACGATGCTCCAATCAAGCAATTTGTATGAGGACGG 300
;
; QY 301 GACATGATCATGACACCCCGGGTGGCTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
; DB 301 GACATGATCATGACACCCCGGGTGGCTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
;
; QY 361 TGCTGGGTAGCGCTCACCCCGGCTAGGAAAGCCAGCATGCCCACTACAAACA 420
; DB 361 TGCTGGGTAGCGCTCACCCCGGCTAGGAAAGCCAGCATGCCCACTACAAACA 420
;
; QY 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGGCTGCTTCTTCTGCTGCTATGCTG 480
; DB 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGGCTGCTTCTTCTGCTGCTATGCTG 480
;
; QY 481 GGGGATCTCTGGGATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
; DB 481 GGGGATCTCTGGGATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
;
;
; US-08-927-597-13
;
; Query Match      100.0%; Score 636; DB 3; Length 636;
; Best Local Similarity 100.0%; Pred. No. 3.3e-164;
; Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 ATGCTGGGTAAGCCATCGATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATT 60
; DB 1 ATGCTGGGTAAGCCATCGATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATT 60
;
; QY 61 CCGCTCGTGGGCGCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120
; DB 61 CCGCTCGTGGGCGCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120
;
; QY 121 CTGGAAGACGGGTGAACCTATGCAACAGGGAATTTGCCCTGCTTCTTCTATCTTC 180
; DB 121 CTGGAAGACGGGTGAACCTATGCAACAGGGAATTTGCCCTGCTTCTTCTATCTTC 180
;
; QY 181 CTCTTGGCTTTACTGCTCTTAACTTCCAGCTTCCGCTTACGAGTGGCGCAACGTG 240
; DB 181 CTCTTGGCTTTACTGCTCTTAACTTCCAGCTTCCGCTTACGAGTGGCGCAACGTG 240
;
; QY 241 TCCGGGATGTACCATGTCAAGAACGATGCTCCAATCAAGCAATTTGTATGAGGACGG 300
; DB 241 TCCGGGATGTACCATGTCAAGAACGATGCTCCAATCAAGCAATTTGTATGAGGACGG 300
;
; QY 301 GACATGATCATGACACCCCGGGTGGCTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
; DB 301 GACATGATCATGACACCCCGGGTGGCTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
;
; QY 361 TGCTGGGTAGCGCTCACCCCGGCTAGGAAAGCCAGCATGCCCACTACAAACA 420
; DB 361 TGCTGGGTAGCGCTCACCCCGGCTAGGAAAGCCAGCATGCCCACTACAAACA 420
;
; QY 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGGCTGCTTCTTCTGCTGCTATGCTG 480
; DB 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGGCTGCTTCTTCTGCTGCTATGCTG 480
;
; QY 481 GGGGATCTCTGGGATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
; DB 481 GGGGATCTCTGGGATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
;
; QY 541 CATGAGACGGTGCAGGATGCAATTTGCTCAATCTATCCCGGCAATAAGGGTACCGGT 600
; DB 541 CATGAGACGGTGCAGGATGCAATTTGCTCAATCTATCCCGGCAATAAGGGTACCGGT 600
;
; QY 601 ATGGCTTGGGATATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 636
; DB 601 ATGGCTTGGGATATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 636
;
;
; RESULT 2
; US-08-927-597-13
; Sequence 13, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```


APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYS, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
NAME/KEY: mat_peptide
LOCATION: 1..789
US-08-612-973-5

Query Match 95.0%; Score 604; DB 3; Length 795;
Best Local Similarity 97.6%; Pred. No. 1.9e-155;
Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1 ATGCTGGGTAAGCCATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGTACATT 60
1 ATGTTGGGTAAGGTATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCATGGGTACATT 60
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
181 CTCTTGGCTTTTACTGCTGCTTAACCATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240
181 CTCTTGGCTTTTACTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTTG 240
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360

US-08-612-973-5

Query Match 96.4%; Score 612.8; DB 3; Length 633;
Best Local Similarity 98.1%; Pred. No. 7e-158;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

1 ATGCTGGGTAAGCCATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGTACATT 60
1 ATGTTGGGTAAGGTATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCATGGGTACATT 60
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
181 CTCTTGGCTTTTACTGCTGCTTAACCATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240
181 CTCTTGGCTTTTACTGCTGCTGCTACCATTTCCAGCTTCCGCTTATGAGGTGCGCAACGTTG 240
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
361 TGCTGGGTAGCGCTCACCCCGACGCTCGCGGTAGGAACGCGCAGCATCTCCCACTACACA 420
361 TGCTGGGTAGCGCTCACCCCGACGCTCGCGGTAGGAACGCGCAGCATCTCCCACTACACA 420
421 ATACGACGCGCATGCGATTGCTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGATGCTG 480
421 ATACGACGCGCATGCGATTGCTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGATGCTG 480
481 GGGGATCTCTGGGATCTGCTTCTCTCTCCAGCTGTTCCAGCTCTCGCTCGCGG 540
481 GGGGATCTCTGGGATCTGCTTCTCTCTCCAGCTGTTCCAGCTCTCGCTCGCGG 540
541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTTACCGT 600
541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTTACCGT 600
601 ATGGCTTGGGATATGATGATGAACCTGTTACTA 632
601 ATGGCTTGGGATATGATGATGAACCTGTTACTA 632

US-08-612-973-5

Query Match 95.0%; Score 604; DB 3; Length 795;
Best Local Similarity 97.6%; Pred. No. 1.9e-155;
Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1 ATGCTGGGTAAGCCATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGTACATT 60
1 ATGTTGGGTAAGGTATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCATGGGTACATT 60
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
181 CTCTTGGCTTTTACTGCTGCTTAACCATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240
181 CTCTTGGCTTTTACTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTTG 240
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
361 TGCTGGGTAGCGCTCACCCCGACGCTCGCGGTAGGAACGCGCAGCATCTCCCACTACACA 420
361 TGCTGGGTAGCGCTCACCCCGACGCTCGCGGTAGGAACGCGCAGCATCTCCCACTACACA 420
421 ATACGACGCGCATGCGATTGCTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGATGCTG 480
421 ATACGACGCGCATGCGATTGCTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGATGCTG 480
481 GGGGATCTCTGGGATCTGCTTCTCTCTCCAGCTGTTCCAGCTCTCGCTCGCGG 540
481 GGGGATCTCTGGGATCTGCTTCTCTCTCCAGCTGTTCCAGCTCTCGCTCGCGG 540
541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTTACCGT 600
541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTTACCGT 600
601 ATGGCTTGGGATATGATGATGAACCTGTTACTA 632
601 ATGGCTTGGGATATGATGATGAACCTGTTACTA 632

US-08-612-973-5

Query Match 95.0%; Score 604; DB 3; Length 795;
Best Local Similarity 97.6%; Pred. No. 1.9e-155;
Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1 ATGCTGGGTAAGCCATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGTACATT 60
1 ATGTTGGGTAAGGTATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCATGGGTACATT 60
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
181 CTCTTGGCTTTTACTGCTGCTTAACCATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240
181 CTCTTGGCTTTTACTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTTG 240
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360

US-08-612-973-5

Query Match 95.0%; Score 604; DB 3; Length 795;
Best Local Similarity 97.6%; Pred. No. 1.9e-155;
Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1 ATGCTGGGTAAGCCATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGTACATT 60
1 ATGTTGGGTAAGGTATCATGATACCCCTTACGTCGGGCTTCGCCGACCTCATGGGTACATT 60
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
61 CCGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGCGCTCGGCGCATGGCGTCCGGGTT 120
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
121 CTGGAAGACGGCGTGAACCTATGCAACAGGGAATTTGCCGTTGCTCTTCTCTATCTTC 180
181 CTCTTGGCTTTTACTGCTGCTTAACCATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240
181 CTCTTGGCTTTTACTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTTG 240
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
241 TCCGGGATGTACCATGTGTCAGGAACGAGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360

US-08-612-973-5

US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:

Db 301 GACATGATCATGCACACCCCGGGTGGTGGTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
QY 361 TGTCTGGGTAGCGCTACCCCGCCAGCTCGCGGGTAGGAACGCCAGCATFCCCACTACAACA 420
Db 361 TGTCTGGGTAGCGCTACCCCGCCAGCTCGCGGGTAGGAACGCCAGCATFCCCACTACAACA 420
QY 421 ATACGACGCGCCAGCTCGATTGCTCGTTGGGGCGGCTGCTTCTGTTCCCGCTATGTACGTG 480
Db 421 ATACGACGCGCCAGCTCGATTGCTCGTTGGGGCGGCTGCTTCTGTTCCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGCGGATCTGCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
Db 481 GGGGATCTCTGCGGATCTGCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
QY 541 CATGACAGCGTGACGACTGCAATGCTCAATCTATCCCGGCCACATACAGGGGTCAACGT 600
Db 541 CATGACAGCGTGACGACTGCAATGCTCAATCTATCCCGGCCACATACAGGGGTCAACGT 600
QY 601 ATGGCTTGGGATATGATGAAGTGGT 628
Db 601 ATGGCTTGGGATATGATGAAGTGGT 628

RESULT 6

US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792

; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789
; US-08-927-597-5

Query Match 95.0%; Score 604; DB 3; Length 795;
Best Local Similarity 97.6%; Pred. No. 1.9e-155;
Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAAGCCATCGATACCTTACGTTGCGGGTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGCTGGGTAAAGCCATCGATACCTTACGTTGCGGGTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CGCTCGTGGCGCCCGCTAGGGGCGCTGCGAGGGCCCTGGCGCATGCGCGTCCGGGTT 120
Db 61 CGCTCGTGGCGCCCGCTAGGGGCGCTGCGAGGGCCCTGGCGCATGCGCGTCCGGGTT 120
QY 121 CTGGAAGAGCGGCGTGAACATGATCAACAGGGAAATTCGCCGGTTCCTTCTATCTTC 180
Db 121 CTGGAAGAGCGGCGTGAACATGATCAACAGGGAAATTCGCCGGTTCCTTCTATCTTC 180
QY 181 CTCTGGCTTTACTGTCCTGTCTAAACATTTCAGAGTTCCGGCTTACGAGGTGCGCAACGTG 240
Db 181 CTCTGGCTTTACTGTCCTGTCTAAACATTTCAGAGTTCCGGCTTACGAGGTGCGCAACGTG 240
QY 241 TCCGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
Db 241 TCCGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATCATGCACACCCCGGGTGGCTGCGCTGCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGCTGCGCTGCGGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGCTCACCCCGCCAGCTCGAGGCTAGGAAGCCAGCGTCCCAACGACA 420
Db 361 TGTGGGTAGCGCTCACCCCGCCAGCTCGAGGCTAGGAAGCCAGCGTCCCAACGACA 420
QY 421 ATACGACGCCAGCTCGATTGCTGCTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCAGCTCGATTGCTGCTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGGGATCTGCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
Db 481 GGGGATCTCTGGGATCTGCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
QY 541 CATGACAGCGTGACGACTGCAATGCTCAATCTATCCCGGCCACATACAGGGGTCAACGT 600
Db 541 CATGACAGCGTGACGACTGCAATGCTCAATCTATCCCGGCCACATACAGGGGTCAACGT 600
QY 601 ATGGCTTGGGATATGATGAAGTGGT 628
Db 601 ATGGCTTGGGATATGATGAAGTGGT 628

RESULT 7

US-08-612-973-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

QY 545 AGACGGTGCAGGACTGCAATTCCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGG 604
Db |||||
QY 545 AGACGGTGCAGGACTGCAATTCCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGG 604
Db |||||
QY 605 CTTGGGATATGATGATGAAGTGGT 628
Db |||||
QY 605 CTTGGGATATGATGATGAAGTGGT 628
Db |||||

RESULT 8
US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..2076
US-08-927-597-47

Query Match 94.6%; Score 601.6; DB 3; Length 2082;
Best Local Similarity 97.8%; Pred. No. 1.1e-154;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 TGGGTAAGGGCCATCGATACCCCTTACGTGCGGCTTCCGCGACCTCGTGGGTACATTCGCG 64
Db |||||
QY 5 TGGGTAAGGTATCATGATACCTTACATGCGGCTTCCGCGACCTCGTGGGTACATTCGCG 64
Db |||||
QY 65 TCCTCGGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGG 124

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..2076
US-08-612-973-47

Query Match 94.6%; Score 601.6; DB 3; Length 2082;
Best Local Similarity 97.8%; Pred. No. 1.1e-154;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 TGGGTAAGGGCCATCGATACCCCTTACGTGCGGCTTCCGCGACCTCGTGGGTACATTCGCG 64
Db |||||
QY 5 TGGGTAAGGTATCATGATACCTTACATGCGGCTTCCGCGACCTCGTGGGTACATTCGCG 64
Db |||||
QY 65 TCGTCGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGG 124
Db |||||
QY 65 TCGTCGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGG 124
Db |||||
QY 125 AAGACGGGTGAATCATGCAACAGGGAATTTGCCGTGGTGTCTTCTTCTATCTTCTCT 184
Db |||||
QY 125 AAGACGGGTGAATCATGCAACAGGGAATTTGCCGTGGTGTCTTCTTCTATCTTCTCT 184
Db |||||
QY 185 TGGCTTTACTGTCTGTCTAACCATTCCAGCTTCCGCTTACGAGTGCGCAACGTGTCGG 244
Db |||||
QY 185 TGGCTTTACTGTCTGTCTAACCATTCCAGCTTCCGCTTACGAGTGCGCAACGTGTCGG 244
Db |||||
QY 245 GGATGTACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTGTATGAGGACGGGACA 304
Db |||||
QY 245 GGATGTACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTGTATGAGGACGGGACA 304
Db |||||
QY 305 TGATATGACACCCCGGGTGGTCCCTGCTTCCGCTTCCGCTTCCGCTTCCGCTGCT 364
Db |||||
QY 305 TGATATGACACCCCGGGTGGTCCCTGCTTCCGCTTCCGCTTCCGCTTCCGCTGCT 364
Db |||||
QY 365 GGGTAGGCTTACCCCGGCTGCTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTGCT 424
Db |||||
QY 365 GGGTAGGCTTACCCCGGCTGCTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTGCT 424
Db |||||
QY 425 GAGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTGGGG 484
Db |||||
QY 425 GAGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTGGGG 484
Db |||||
QY 485 ATCTCTGGGATCTGCTTCTCTCCAGCTGTTTCCAGCTGTTTCCAGCTGTTTCCAGCTGTT 544
Db |||||
QY 485 ACCTCTCGGATCTGCTTCTCTCCAGCTGTTTCCAGCTGTTTCCAGCTGTTTCCAGCTGTT 544
Db |||||

Db 65 TCGTGGGCGCCCTAGGGGGCGCTGCACGGCCCTGGCGCATGGCTCCGGGTTCTGG 124
QY 125 AGACGGCGTGAATATGCAACAGGAAATTTGCTGCTTTCTCTATCTTCTCTCT 184
Db 125 AGACGGCGTGAATATGCAACAGGAAATTTGCTGCTTTCTCTATCTTCTCTCT 184
QY 185 TGGCTTTACTGCTGCTCTAAACATTCAGCTTCGCTTACGAGGTGCGCAAGTTCGG 244
Db 185 TGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGCTTATGAAGTGCAGCAAGTTCGG 244
QY 245 GGATGTACCATTGTACGAAACGACTGTCTCAACACTCAAGCAATTTGTATGAGGACGCGACA 304
Db 245 GGATGTACCATTGTACGAAACGACTGTCTCAACACTCAAGCAATTTGTATGAGGACGCGACA 304
QY 305 TGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCGTTCGGGAGAACAACTCTTCCCGTGTCT 364
Db 305 TGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCGTTCGGGAGAACAACTCTTCCCGTGTCT 364
QY 365 GGGTAGCGGTCAACCCCGCGCTGCGGCTAGGACGCGACGATCCCGCACTACAACTATAC 424
Db 365 GGGTAGCGGTCAACCCCGCGCTGCGGCTAGGACGCGACGATCCCGCACTACAACTATAC 424
QY 425 GACGCCACGTGCAATTTGCTGCTGGGGCGGCTGCTTTCTGCTATGTAGTGGGG 484
Db 425 GACGCCACGTGCAATTTGCTGCTGGGGCGGCTGCTTTCTGCTATGTAGTGGGG 484
QY 485 ATCTCTGCGGATGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
Db 485 ACCTCTGCGGATGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
QY 545 AGACGGTGCAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
Db 545 AGACGGTGCAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
QY 605 CTGGGATATGATGATGAATGCT 628
Db 605 CTGGGATATGATGATGAATGCT 628

RESULT 9

US-08-612-973-49, Application US/08612973
; Sequence 49, Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUISE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2430
; NAME/KEY: mat_peptide
; LOCATION: 1..2427
; US-08-612-973-49

Query Match 94.6%; Score 601.6; DB 3; Length 2433;
Best Local Similarity 97.8%; Pred. No. 1.2e-154;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 TGGGTAAGGCCATCGATACCCCTAGCGGGCGCTGCGAGGGCCCTGCGCGATGCGCTCCGGGTTCTGG 64
Db 356 TGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGCG 415
QY 65 TCGTGGCGCCCGCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGCTCCGGGTTCTGG 124
Db 416 TCGTGGCGCCCGCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGCTCCGGGTTCTGG 475
QY 125 AAGACGGCGTGAATATGCAACAGGGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 184
Db 476 AAGACGGCGTGAATATGCAACAGGGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535
QY 185 TGGCTTTACTGCTGCTGCTTAAACATTCAGCTTCGCTTACGAGGTGCGCAAGCTGCTCG 244
Db 536 TGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
QY 245 GGATGTACCATTGTACGAAACGACTGCTCCAACCTCAAGCAATTTGTATGAGGACGCGACA 304
Db 596 GGATGTACCATTGTACGAAACGACTGCTCCAACCTCAAGCAATTTGTATGAGGACGCGACA 655
QY 305 TGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
Db 656 TGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
QY 365 GGGTAGCGGTCAACCCCGCGCTGCGGCTAGGAAACGCGAGCATCCCGCACTACAACTATAC 424
Db 716 GGGTAGCGGTCAACCCCGCGCTGCGGCTAGGAAACGCGAGCATCCCGCACTACAACTATAC 775
QY 425 GACGCCACGTGCAATTTGCTGCTGGGGCGGCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484
Db 776 GACGCCACGTGCAATTTGCTGCTGGGGCGGCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
QY 485 ATCTCTGCGGATGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
Db 836 ACCTCTGCGGATGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGGCCACATAACGGGTACCGGTATGG 604
Db 896 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGGCCACATAACGGGTACCGGTATGG 955
QY 605 CTGGGATATGATGATGAATGCT 628
Db 956 CTGGGATATGATGATGAATGCT 979

RESULT 10

US-08-927-597-49
; Sequence 49, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2430
NAME/KEY: mat_peptide
LOCATION: 1..2427

US-08-927-597-49
Query Match 94.6%; Score 601.6; DB 3; Length 2433;
Best Local Similarity 97.8%; Pred. No. 1.2e-154;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
5 TGGGTAAGCCATCGATACCTTACGTGGCGCTTCGCCGACCTCGTGGGGTACATTCCGC 64
356 TGGGTAAGGTCATCGATACCTTACGTGGCGCTTCGCCGACCTCGTGGGGTACATTCCGC 415
65 TCCTCGCGCGCCCGCCCTAGGCGGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGGTTCTGG 124
416 TCCTCGCGCGCCCGCCCTAGGCGGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGGTTCTGG 475
125 AAGACGGCGTGAATCATGCAAGGGAATTCGCTGTTGCTCTTCTATCTTCTCTCT 184
476 AGGACGGCGTGAATCATGCAAGGGAATTCGCTGTTGCTCTTCTATCTTCTCTCT 535
185 TGGCTTTACTGCTCTGCTTAACCAATTCAGCTTCCGCTTACGAGGTGCGCAACGTTCCG 244
536 TGGCTTTGCTGCTCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGCACACGTTCCG 595
245 GGATGTACCATGTGACCAAGGCTGCTCCAACTCAAGCATTTGTATGAGGCGCGGACA 304
596 GGATGTACCATGTGACCAAGGCTGCTCCAACTCAAGCATTTGTATGAGGCGCGGACA 655

305 TGATCATGCACACCCCGGGTGGTGGCTGGGAGAAACAATCTTCCCGTGTCT 364
656 TGATCATGCACACCCCGGGTGGTGGCTGGGAGAAACAATCTTCCCGTGTCT 715
365 GGGTAGCGCTCACCCCGACGCTCGCGGCTAGGAAGCCAGCATGCCCAACAATAC 424
716 GGGTAGCGCTCACCCCGACGCTCGCGGCTAGGAAGCCAGCATGCCCAACAATAC 775
425 GAGCCACGCTCGAATTTGCTGCTGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 484
776 GAGCCACGCTCGAATTTGCTGCTGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 835
485 ATCTCTCGGATCTGCTTCTCTGCTTCCAGCTGTTTCCAGCATCTCGCTCGCGGCATG 544
836 ACTCTCGGATCTGCTTCTCTGCTTCCAGCTGTTTCCAGCATCTCGCTCGCGGCATG 895
545 AGCGGTGCAGGACTGCAATTTGCTGCTTATCCCGCCACATACGGGTCAACGATGG 604
896 AGCGGTGCAGGACTGCAATTTGCTGCTTATCCCGCCACATACGGGTCAACGATGG 955
605 CTTGGGATATGATGATGAATGTT 628
956 CTTGGGATATGATGATGAATGTT 979

RESULT 11
US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-470-426B-17

Query Match 87.0%; Score 553.6; DB 2; Length 1539;
Best Local Similarity 92.9%; Pred. No. 1.2e-141;
Matches 580; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 5 TGGGTAAGGCCATCGATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGGTACATTCGCCG 64
DB TGGGTAAGGTCATCGATACCCCTTACATGCGCTTCGCCGATCTCATGGGGTATATTCGCC 415
QY 65 TCGTGGCGCCGCCCTAGGGGGCGTGCAGAGCCCTGGGGGATGGCGTCCGGGTTCTGG 124
DB TCGTGGCGCCGCCCTAGGGGGCGTGCAGAGCCCTGGGGGATGGCGTCCGGGTTCTGG 475
QY 125 AAGACGGCTGAACATGCAACAGGGAATTCGCTGTTGCTTCTCTATCTCTCTCT 184
DB AGACGGCGTGAACATGCAACAGGGAATTCGCTGTTGCTTCTCTATCTCTCTCTCT 535
QY 185 TGGCTTTACTGCTCTCTAACCATTCCAGCTTCGCCCTTACGAGGTGCGCAACGTGCG 244
DB TGGCTTTACTGCTCTCTAACCATTCCAGCTTCGCCCTTACGAGGTGCGCAACGTGCG 595
QY 245 GGATGTACCATGTACGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCGGACA 304
DB GGATATACCATGTACGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCGGACA 655
QY 305 TGATCATGCACACCCCGGGTGGTGGCTCGGTTCCGGGAGAACAACTCTTCCCGTCTG 364
DB TGATCATGCATCTCCCGGGTGGTGGCTCGGTTCCGGGAGAACAACTCTTCCCGTCTG 715
QY 365 GGGTAGCGCTACCCCGGCTCGCGGCTAGGAACGCCAGCATCCCGCACTACAAACAATAC 424
DB GGGTAGCGCTACTCCCGGCTCGCGGCTAGGAACGCCAGCATCCCGCACTACAAACAATAC 775
QY 425 GACGCCACGTGCAATTTGCTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTTGGGG 484
DB GACGCCACGTGCAATTTGCTTGGGGCGGCTGCTTTCTGCTCGCTATGTACGTTGGGG 835
QY 485 ATCTCTGCGGATGCTCTCTGCTCCGCTCCAGCTGTTCCACCTCTCGCTCGCGGCATG 544
DB ATCTCTGCGGATGCTCTCTGCTCCGCTCCAGCTGTTCCACCTCTCGCTCGCGGCATG 895
QY 545 AGACGGTGAGGACTGCAATTTGCTCAATCTATCCCGGCACATAACGGGTACCGGTATGG 604
DB AGACGGTGAGGACTGCAATTTGCTCAATCTATCCCGGCACATAATTCAGGTACCGGTATGG 955
QY 605 CTTGGGATATGATGAATGCT 628
DB CTTGGGATATGATGAATGCT 979

RESULT 12

US-08-470-426B-14
Sequence 14, Application US/08470426B
Patent No. 5856458
GENERAL INFORMATION:
APPLICANT: Okamoto, Hiroaki
APPLICANT: Nakamura, Tetsuo
TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
TITLE OF INVENTION: HEPATITIS VIRUS
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young,
ADDRESSEE: L.L.P.
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,426B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-153402
FILING DATE: 12-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Wellacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/59-47083.1
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-470-426B-14

Query Match 87.0%; Score 553.6; DB 2; Length 1863;
Best Local Similarity 92.9%; Pred. No. 1.3e-141;
Matches 580; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 5 TGGGTAAGGCCATCGATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGGTACATTCGCCG 64
DB TGGGTAAGGTCATCGATACCCCTTACATGCGCTTCGCCGATCTCATGGGGTATATTCGCC 739
QY 65 TCGTGGCGCCGCCCTAGGGGGCGTGCAGAGCCCTGGCGCATGGCGTCCGGGTTCTGG 124
DB TCGTGGCGCCGCCCTAGGGGGCGTGCAGAGCCCTGGCGCATGGCGTCCGGGTTCTGG 799
QY 125 AAGACGGCTGAACATGCAACAGGGAATTCGCTGTTGCTTCTCTATCTCTCTCT 184
DB AGACGGCGTGAACATGCAACAGGGAATTCGCTGTTGCTTCTCTATCTCTCTCTCT 859
QY 185 TGGCTTTACTGCTCTCTAACCATTCCAGCTTCGCCCTTACGAGGTGCGCAACGTGCG 244
DB TGGCTTTACTGCTCTCTAACCATTCCAGCTTCGCCCTTACGAGGTGCGCAACGTGCG 919
QY 245 GGATGTACCATGTACGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCGGACA 304
DB GGATATACCATGTACGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCGGACA 979
QY 305 TGATCATGCACACCCCGGGTGGTGGCTCGGTTCCGGGAGAACAACTCTTCCCGTCTG 364
DB TGATCATGCATCTCCCGGGTGGTGGCTCGGTTCCGGGAGAACAACTCTTCCCGTCTG 1039
QY 365 GGGTAGCGCTACCCCGGCTCGCGGCTAGGAACGCCAGCATCCCGCACTACAAACAATAC 424
DB GGGTAGCGCTCACTCCCGGCTCGCGGCTAGGAACGCCAGCATCCCGCACTACAAACAATAC 1099
QY 425 GACGCCACGTGCAATTTGCTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTTGGGG 484
DB GACGCCACGTGCAATTTGCTTGGGGCGGCTGCTTTCTGCTCGCTATGTACGTTGGGG 1159
QY 485 ATCTCTGCGGATGCTCTCTGCTCCGCTCCAGCTGTTCCACCTCTCGCTCGCGGCATG 544
DB ATCTCTGCGGATGCTCTCTGCTCCGCTCCAGCTGTTCCACCTCTCGCTCGCGGCATG 1219
QY 545 AGACGGTGAGGACTGCAATTTGCTCAATCTATCCCGGCACATAACGGGTACCGGTATGG 604
DB AGACGGTGAGGACTGCAATTTGCTCAATCTATCCCGGCACATAATTCAGGTACCGGTATGG 1279
QY 605 CTTGGGATATGATGAATGCT 628
DB CTTGGGATATGATGAATGCT 1303

RESULT 13

[illegible]

Db 190 TGGCTTTGCTGCTCTTTGACCATCCAGCTTCCGCTTATGAGTGCGCAAGCATATCCG 249
QY 245 GGATGTACATGTACGAACGACTGTCTCAACTCAAGCATTTGTATGAGCGAGCGGACA 304
Db 250 GGATATACATGTACGAACGACTGTCTCAACTCAAGCATTTGTATGAGCGAGCGGACA 309
QY 305 TGATCATGCACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
Db 310 TGATCATGCATACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
QY 365 GGATAGCTGTACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
Db 370 GGCGAGCTGTACTCCAGCTGTAGCGCGGAGGACACCGGCTCCCACTACGACATATC 429
QY 425 GACGCCAGCTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484
Db 430 GACGGCATGTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
QY 485 ATCTCTGCGATCTGCT 544
Db 490 ATCTCTGCGATCTGCT 549
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAAGCGGTACCGGTATGG 604
Db 550 AGACAGTACAGGACTGCAACTGCTCAATCTATCCCGGCGCACTTGACAGGTATCGCATGG 609
QY 605 CTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
Db 610 CTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 633

RESULT 15

US-08-462-195-1
; Sequence 1, Application US/08462195
; Patent No. 5789544
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: MATSUURA, YOSHIHARU
; APPLICANT: HONDA, YOSHIKAZU
; APPLICANT: SEKI, MAKOTO
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
; TITLE OF INVENTION: HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,195
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,303
; FILING DATE: 22-MAY-1995
; APPLICATION NUMBER: US 08/074,584
; FILING DATE: 11-JUN-1993
; APPLICATION NUMBER: JP 152487/1992
; FILING DATE: 11-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5789544man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4169-003-0
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855.OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1037 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C virus
; IMMEDIATE SOURCE:
; CLONE: pUC010
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 17..1036
; US-08-462-195-1

Query Match 86.3%; Score 548.8; DB 1; Length 1037;
Best Local Similarity 92.5%; Pred. No. 2.2e-140;
Matches 577; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 5 TGGGTAAGCCCATCGATACCCCTTACGTGCGGCTTTCGCCGACCTCGTGGGGGTACATTCGCG 64
Db 372 TGGGTAAGGTATCGATACCCCTTACATGCGGCTTTCGCCGACCTCATGCGGGTACATCCCGC 431
QY 65 TCGTGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGCGGATGCGGTCGGGTTCTGG 124
Db 432 TTGTGCGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGCGACATGTTGTGTCGGGTTCTGG 491
QY 125 AAGACGGCTGAACTATGCAACAGGGAATTTGCTGTTGCTTTCTCTCTATCTCTCTCT 184
Db 492 AGGACGGCTGAACTATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTCTCTCT 551
QY 185 TGGCTTTACTGCTGCTTAACCAATTCAGCTTCGCTTACGAGGTGCGCAACGTGTCCG 244
Db 552 TAGCTCTGCTGCTTTTGGACCATCCAGCTTCCGCTTATGAAGTGCAGCAACGTGTCCG 611
QY 245 GGATGTACCATGTACGAACGACTGCTCAACTCAAGCATTTGTATGAGCGAGCGGACA 304
Db 612 GGATATACCATGTACAAACGACTGCTCCAATCAAGCATTTGTATGAGCGGCGGAGCG 671
QY 305 TGATCATGCACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
Db 672 TGATCATGCATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
QY 365 GGGTAGCGCTACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
Db 732 GGGTAGCGCTCACTCCAGCTGCGGGCGAGGAATGCCAGCGTCCCACTACGACATTAC 791
QY 425 GAGGCCAGCTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484
Db 792 GAGGCCAGCTCGACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
QY 485 ATCTCTGCGGATCTGCTTCT 544
Db 852 ATCTCTGCGGATCTGTTTCT 911
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAAGCGGTACCGGTATGG 604
Db 912 AGACAGTACAGGACTGCAACTGCTCAATCTATCCCGGCGCACGTATCAGGCCATCGTATGG 971
QY 605 CTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
Db 972 CTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 995

Search completed: December 20, 2003, 07:03:03
Job time : 47.8661 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2737.48 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-21

Perfect score: 723

Sequence: 1 ATGTTGGTAGGTCATCGA.....TACTCTTGTCTCCCTAATAG 723

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pln.*

35: em.htg_rod.*

36: em.htg_mam.*

37: em.htg_vrt.*

38: em.sy.*

39: em.htgo_hum.*

40: em.htgo_mus.*

41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	723	100.0	723	6	A48683	A48683 Sequence 21
2	723	100.0	723	6	AR157336	AR157336 Sequence
3	723	100.0	723	6	AX452770	AX452770 Sequence
4	723	100.0	723	6	AX685022	AX685022 Sequence
5	641	88.7	795	6	A48667	A48667 Sequence 5
6	641	88.7	795	6	AR157325	AR157325 Sequence
7	641	88.7	795	6	AX452754	AX452754 Sequence
8	641	88.7	795	6	AX685006	AX685006 Sequence
9	624.8	86.4	2082	6	A48709	A48709 Sequence 47
10	624.8	86.4	2082	6	AR157350	AR157350 Sequence
11	624.8	86.4	2082	6	AX452796	AX452796 Sequence
12	624.8	86.4	2082	6	AX685048	AX685048 Sequence
13	624.8	86.4	2433	6	A48711	A48711 Sequence 49
14	624.8	86.4	2433	6	AR157351	AR157351 Sequence
15	624.8	86.4	2433	6	AX452798	AX452798 Sequence
16	624.8	86.4	2433	6	AX685050	AX685050 Sequence
17	598.2	82.7	606	6	A48687	A48687 Sequence 25
18	598.2	82.7	606	6	AR157338	AR157338 Sequence
19	598.2	82.7	606	6	AX452774	AX452774 Sequence
20	598.2	82.7	606	6	AX685026	AX685026 Sequence
21	597	82.6	636	6	A48689	A48689 Sequence 27
22	597	82.6	636	6	AR157339	AR157339 Sequence
23	597	82.6	636	6	AX452776	AX452776 Sequence
24	597	82.6	636	6	AX685028	AX685028 Sequence
25	565.6	78.2	3296	14	AB008441	AB008441 Hepatitis
26	565.6	78.2	3296	14	AB008443	AB008443 Hepatitis
27	565	78.1	9418	14	HCV132996	HCV132996 Hepatitis
28	564	78.0	3296	14	AB008442	AB008442 Hepatitis
29	564	78.0	9379	14	AF207766	AF207766 Hepatitis
30	564	78.0	9379	14	HCVPOLYP	HCVPOLYP Hepatitis
31	562.4	77.8	9386	14	AF165056	AF165056 Hepatitis
32	560.8	77.6	3296	14	AB008446	AB008446 Hepatitis
33	560.8	77.6	9386	14	AF165055	AF165055 Hepatitis
34	559.2	77.3	3296	14	AB008447	AB008447 Hepatitis
35	559.2	77.3	9379	14	AF165052	AF165052 Hepatitis
36	559.2	77.3	9379	14	AF207761	AF207761 Hepatitis
37	559.2	77.3	9410	14	HPCKIR2	D50481 Hepatitis C
38	559.2	77.3	9605	6	AX739971	AX739971 Sequence
39	559.2	77.3	9605	14	HCV238799	HCV238799 Hepatitis
40	559.2	77.3	11076	6	AX036252	AX036252 Sequence
41	559.2	77.3	11076	6	AX036258	AX036258 Sequence
42	559.2	77.3	11076	6	AX036260	AX036260 Sequence
43	559.2	77.3	11076	6	AX036262	AX036262 Sequence
44	557.6	77.1	3296	14	AB008445	AB008445 Hepatitis
45	557.6	77.1	9033	14	HCV238800	HCV238800 Hepatitis

ALIGNMENTS

RESULT 1
A48683
LOCUS
Sequence 21 from Patent WO9604385.
DEFINITION
A48683
ACCESSION
A48683.1 GI:2302396
VERSION
KEYWORDS
SOURCE
ORGANISM
unidentified
unclassified.
REFERENCE
1 (bases 1 to 723)
Maertens,G., Bosman,F., De.M.G. and Buyse,M.
TITLE
PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL
Patent: WO 9604385-A 21 15-FEB-1996;

linear PAT 07-MAR-1997

[illegible]


```
Db 601 CAAGCTGTGTGACATGTTGGCGGGGCCCATTTGGGGAGTCTCGGGGTCTCGCCTAC 660
QY 661 TATTCATCGTGGGAACTGGGCTAAGGTTTGAATGTGATGCTACTCTTTTGTCTCCCTAA 720
Db 661 TATTCATCGTGGGAACTGGGCTAAGGTTTGAATGTGATGCTACTCTTTTGTCTCCCTAA 720
QY 721 TAG 723
Db 721 TAG 723

RESULT 3
AX452770
LOCUS AX452770 723 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 21 from Patent EP1211315.
ACCESSION AX452770
VERSION AX452770.1 GI:21712455
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; SERNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 21 05-JUN-2002;
INNOGENETICS N.V. (BE)
FEATURES
Location/Qualifiers
source 1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38072.1"
/db_xref="GI:21712456"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGAAARALAHGVRLVDG
VNYATGNLPGCSFIFLLALLSLTPASAYEVNRVSGMYHTVNDSCNSSIIVEAADM
IMHTGCVPCVRNNSRCWALPTLAARNASVPTTIRRHVDLSQFTISPRRHETV
ODNCNSIYPGHITGHRMADMMNWSPTTALVVSQLLRIPQAVDVNVAGAHGVLGL
AYSVNGWAKVLIIVLLFAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
ORIGIN

Query Match 100.0%; Score 723; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 7.4e-148;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGTAAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGTAAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATCGCGTCCGGTT 120
Db 61 CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATCGCGTCCGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTCCCGGGTGTCTTTCTCTAFTTC 180
Db 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTCCCGGGTGTCTTTCTCTAFTTC 180
QY 181 CTCCTGGCTTCTGTCCTGTCTGACCGTTCAGCTTCGCTTATGAAGTGGCAACGTG 240
Db 181 CTCCTGGCTTCTGTCCTGTCTGACCGTTCAGCTTCGCTTATGAAGTGGCAACGTG 240
QY 241 TCGGGATGTACCATGTACAGAACACTGCTCCAACTCAAGCATTCGTATGAGGACGCG 300
Db 241 TCGGGATGTACCATGTACAGAACACTGCTCCAACTCAAGCATTCGTATGAGGACGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGTGCGTCCCTGCGTTTCGGGAGAACAACTCTTCCCGC 360
```

```
Db 301 GACATGATCATGCACACCCCGGGTGTGCGTCCCTGCGTTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGTCTGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAGCCAGCGTCCCCACACGACA 420
Db 361 TGTCTGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAGCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCAAGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 480
Db 421 ATACGACGCAAGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 480
QY 481 CAGGACTGCAATTCGCTCAATCTATCCGGGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTCGCTCAATCTATCCGGGCCACATAACGGGTACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACCTGGTTCGCTACACCGGCCCTGGTGTATCGCAGCTGCTCCGATCCCA 600
Db 541 ATGATGATGAACCTGGTTCGCTACACCGGCCCTGGTGTATCGCAGCTGCTCCGATCCCA 600
QY 601 CAAGCTGTCTGGACATGCTGGCGGGGCCAATTTGGGGAGTCTTGGCGGGTCTCGCCTAC 660
Db 601 CAAGCTGTCTGGACATGCTGGCGGGGCCAATTTGGGGAGTCTTGGCGGGTCTCGCCTAC 660
QY 661 TATTCATCGTGGGAACTGGGCTAAGGTTTGAATGTGATGCTACTCTTTTGTCTCCCTAA 720
Db 661 TATTCATCGTGGGAACTGGGCTAAGGTTTGAATGTGATGCTACTCTTTTGTCTCCCTAA 720
QY 721 TAG 723
Db 721 TAG 723

RESULT 4
AX685022
LOCUS AX685022 723 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 21 from Patent WO0205548.
ACCESSION AX685022
VERSION AX685022.1 GI:29371427
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; SERNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Maertens,G., Bosman,F. and Buyse,M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 21 18-JUL-2002;
INNOGENETICS N.V. (BE)
FEATURES
Location/Qualifiers
source 1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD86525.1"
/db_xref="GI:29371428"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGAAARALAHGVRLVDG
VNYATGNLPGCSFIFLLALLSLTPASAYEVNRVSGMYHTVNDSCNSSIIVEAADM
IMHTGCVPCVRNNSRCWALPTLAARNASVPTTIRRHVDLSQFTISPRRHETV
ODNCNSIYPGHITGHRMADMMNWSPTTALVVSQLLRIPQAVDVNVAGAHGVLGL
AYSVNGWAKVLIIVLLFAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
ORIGIN

Query Match 100.0%; Score 723; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 7.4e-148;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```

/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

1..792
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAA03129.1"
/db_xref="GI:2302381"
/translation="MFLGKVIDTLTCGFADLVGVPLVGA PLGGAARALAHGVRLVDG
VNYATGNLPGCSFILLALLSLTLPASAYENRVNMGVHYVTDNSNSIVFAADM
IMHPCGVPCVRENNSRCRWALPTPLAARNA SVPTTIIIRHVDLLVGLAAAFCSAMV
GDLGCSVFLSQLFTISPRRHETVQDCNCSIYPGHTIGHRMAMWMMNMSPTTALVVS
QLLRIPQAVDMVAGHWGLAGLAYISMVGNMAKVLIVMLFAP"

1..789
/product="unnamed"

mat_peptide
BASE COUNT 130 a 240 c 231 g 194 t
ORIGIN

Query Match 88.7%; Score 641; DB 6; Length 795;
Best Local Similarity 90.9%; Pred. No. 6.3e-130;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 1 ATGTTGGGTAAGGTATCATGATACCTTACATGCGGCTTGC CGGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAGGTATCATGATACCTTACATGCGGCTTGC CGGACCTCGTGGGTACATT 60

QY 61 CCCTCGTGGCGGCCCTTAGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB 61 CCCTCGTGGCGGCCCTTAGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 61 CCCTCGTGGCGGCCCTTAGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB 61 CCCTCGTGGCGGCCCTTAGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGGAAGA CGGCGTGAACCTATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180
DB 121 CTGGAAGA CGGCGTGAACCTATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180

QY 121 CTGGAAGA CGGCGTGAACCTATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180
DB 121 CTGGAAGA CGGCGTGAACCTATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180

QY 181 CTCTTGGCTTTGTCTGTCTGTCTGAACCGTTTCAGCTTCGCTTATGAAGTGGCAACGTT 240
DB 181 CTCTTGGCTTTGTCTGTCTGTCTGAACCGTTTCAGCTTCGCTTATGAAGTGGCAACGTT 240

QY 241 TCCGGGATGATCAATGTACGAACGACTGCTCCACTCAAGCATTTGTGTATGAGCAGCAGG 300
DB 241 TCCGGGATGATCAATGTACGAACGACTGCTCCACTCAAGCATTTGTGTATGAGCAGCAGG 300

QY 301 GACATGATCATGACACACCCCGGGTGCCTCGCTTGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGCCTCGCTTGGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCAACCCCAACGCTCGAGCTAGGAACGCCAGCGTCCCAACGACA 420
DB 361 TGCTGGGTAGCGCTCAACCCCAACGCTCGAGCTAGGAACGCCAGCGTCCCAACGACA 420

QY 421 ATACGAGCCACAGTTCGAT-----TCCAGCTTCGCTTCTGTTCCGCTATGTACGTG 480
DB 421 ATACGAGCCACAGTTCGAT-----TCCAGCTTCGCTTCTGTTCCGCTATGTACGTG 480

QY 439 -----TCCAGCTTCGCTTCTGTTCCGCTATGTACGTG 468
DB 439 -----TCCAGCTTCGCTTCTGTTCCGCTATGTACGTG 468

QY 481 GGGGACCTCTGCGGATCTGTCTTCTCGTCTCCAGCTGTTTCCATCTCGCTCCGCG 540
DB 481 GGGGACCTCTGCGGATCTGTCTTCTCGTCTCCAGCTGTTTCCATCTCGCTCCGCG 540

QY 469 CATGAGACGCTGACGACTGCAATTTGTCTCAATCTTCCCGCCACATPAAAGGTCACCGT 528
DB 469 CATGAGACGCTGACGACTGCAATTTGTCTCAATCTTCCCGCCACATPAAAGGTCACCGT 528

QY 541 CATGAGACGCTGACGACTGCAATTTGTCTCAATCTTCCCGCCACATPAAAGGTCACCGT 600
DB 541 CATGAGACGCTGACGACTGCAATTTGTCTCAATCTTCCCGCCACATPAAAGGTCACCGT 600

QY 529 ATGGCTTGGGATATGATGAATGTCGCTGCTACACGGCCCTGGTGGTATCGACGTG 588
DB 529 ATGGCTTGGGATATGATGAATGTCGCTGCTACACGGCCCTGGTGGTATCGACGTG 588

QY 601 ATGGCTTGGGATATGATGAATGTCGCTGCTACACGGCCCTGGTGGTATCGACGTG 660
DB 601 ATGGCTTGGGATATGATGAATGTCGCTGCTACACGGCCCTGGTGGTATCGACGTG 660

QY 589 CTCGGGATCCCAACAGCTGTCTGTGACATGTTGGCGGGGCCCATTTGGGAGTCTCTGGCG 648
DB 589 CTCGGGATCCCAACAGCTGTCTGTGACATGTTGGCGGGGCCCATTTGGGAGTCTCTGGCG 648

QY 661 CTCGGGATCCCAACAGCTGTCTGTGACATGTTGGCGGGGCCCATTTGGGAGTCTCTGGCG 708
DB 661 CTCGGGATCCCAACAGCTGTCTGTGACATGTTGGCGGGGCCCATTTGGGAGTCTCTGGCG 708

QY 649 GGTCTCGCTACTATTTCATGTTGGGAACTGGGCTAAGTTTTGATGTGATGCTACTC 708
DB 649 GGTCTCGCTACTATTTCATGTTGGGAACTGGGCTAAGTTTTGATGTGATGCTACTC 708

QY 721 GGTCTCGCTACTATTTCATGTTGGGAACTGGGCTAAGTTTTGATGTGATGCTACTC 788
DB 721 GGTCTCGCTACTATTTCATGTTGGGAACTGGGCTAAGTTTTGATGTGATGCTACTC 788

/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

1..792
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAA03129.1"
/db_xref="GI:2302381"
/translation="MFLGKVIDTLTCGFADLVGVPLVGA PLGGAARALAHGVRLVDG
VNYATGNLPGCSFILLALLSLTLPASAYENRVNMGVHYVTDNSNSIVFAADM
IMHPCGVPCVRENNSRCRWALPTPLAARNA SVPTTIIIRHVDLLVGLAAAFCSAMV
GDLGCSVFLSQLFTISPRRHETVQDCNCSIYPGHTIGHRMAMWMMNMSPTTALVVS
QLLRIPQAVDMVAGHWGLAGLAYISMVGNMAKVLIVMLFAP"

1..789
/product="unnamed"

mat_peptide
BASE COUNT 130 a 240 c 231 g 194 t
ORIGIN

Query Match 88.7%; Score 641; DB 6; Length 795;
Best Local Similarity 90.9%; Pred. No. 6.3e-130;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 1 ATGTTGGGTAAGGTATCATGATACCTTACATGCGGCTTGC CGGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAGGTATCATGATACCTTACATGCGGCTTGC CGGACCTCGTGGGTACATT 60

QY 61 CCCTCGTGGCGGCCCTTAGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB 61 CCCTCGTGGCGGCCCTTAGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 61 CCCTCGTGGCGGCCCTTAGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB 61 CCCTCGTGGCGGCCCTTAGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGGAAGA CGGCGTGAACCTATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180
DB 121 CTGGAAGA CGGCGTGAACCTATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180

QY 121 CTGGAAGA CGGCGTGAACCTATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180
DB 121 CTGGAAGA CGGCGTGAACCTATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180

QY 181 CTCTTGGCTTTGTCTGTCTGTCTGAACCGTTTCAGCTTCGCTTATGAAGTGGCAACGTT 240
DB 181 CTCTTGGCTTTGTCTGTCTGTCTGAACCGTTTCAGCTTCGCTTATGAAGTGGCAACGTT 240

QY 241 TCCGGGATGATCAATGTACGAACGACTGCTCCACTCAAGCATTTGTGTATGAGCAGCAGG 300
DB 241 TCCGGGATGATCAATGTACGAACGACTGCTCCACTCAAGCATTTGTGTATGAGCAGCAGG 300

QY 301 GACATGATCATGACACACCCCGGGTGCCTCGCTTGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGCCTCGCTTGGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCAACCCCAACGCTCGAGCTAGGAACGCCAGCGTCCCAACGACA 420
DB 361 TGCTGGGTAGCGCTCAACCCCAACGCTCGAGCTAGGAACGCCAGCGTCCCAACGACA 420

QY 421 ATACGAGCCACAGTTCGAT-----TCCAGCTTCGCTTCTGTTCCGCTATGTACGTG 480
DB 421 ATACGAGCCACAGTTCGAT-----TCCAGCTTCGCTTCTGTTCCGCTATGTACGTG 480

QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTTCAACCGTATGGCTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTTCAACCGTATGGCTTGGGAT 540

QY 541 ATGATGATGAATGTCGCTTCCAGCTGTTCAACCTGCTCGCTCGCGCATGAGACGGTGT 600
DB 541 ATGATGATGAATGTCGCTTCCAGCTGTTCAACCTGCTCGCTCGCGCATGAGACGGTGT 600

QY 541 ATGATGATGAATGTCGCTTCCAGCTGTTCAACCTGCTCGCTCGCGCATGAGACGGTGT 600
DB 541 ATGATGATGAATGTCGCTTCCAGCTGTTCAACCTGCTCGCTCGCGCATGAGACGGTGT 600

QY 601 CAACTGCTGAGACATGTTGGCGGGGCCCAATTGGGAGTCTTGGCGGCTTCGCGCTAC 660
DB 601 CAACTGCTGAGACATGTTGGCGGGGCCCAATTGGGAGTCTTGGCGGCTTCGCGCTAC 660

QY 601 CAACTGCTGAGACATGTTGGCGGGGCCCAATTGGGAGTCTTGGCGGCTTCGCGCTAC 660
DB 601 CAACTGCTGAGACATGTTGGCGGGGCCCAATTGGGAGTCTTGGCGGCTTCGCGCTAC 660

QY 661 TATTCATGTTGGGAACTGGGCTAAGTTTTGATTTGATGCTACTCTTTGCTCCCTAA 720
DB 661 TATTCATGTTGGGAACTGGGCTAAGTTTTGATTTGATGCTACTCTTTGCTCCCTAA 720

QY 721 TAG 723
DB 721 TAG 723

RESULT 5
LOCUS A48667
DEFINITION Sequence 5 from Patent WO9604385.
ACCESSION A48667
VERSION A48667.1 GI:2302380
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 795)
AUTHORS Maertens G., Boeman, F., De, M.G. and Buyse, M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 5 15-FEB-1996;
INNOGENETICS NV (BE)
COMMENT Other publication AU 2172273 960215
Other publication AU 3382495 960304.
Location/Qualifiers
FEATURES
1..795

```


Db	241	TCGGGATGTACCATGTCTACGAA	CGACTGCTCCAACTCAAGCATTTGTATGAGGCAGCG	300	
Qy	301	GACATGATCATGCAACACCCCGGGTGGTCCCTTCCGGAGAACAACTCTTCCCGC	360		
Db	301	GACATGATCATGCAACACCCCGGGTGGTCCCTTCCGGAGAACAACTCTTCCCGC	360		
Qy	361	TGCTGGGTAGCGCTACACCCACCGTCCGACGTAGGAACGCGAGCGTCCCAACAGACA	420		
Db	361	TGCTGGGTAGCGCTACACCCACCGTCCGACGTAGGAACGCGAGCGTCCCAACAGACA	420		
Qy	421	ATACGACGCCAGTCCGAT-----	438		
Db	421	ATACGACGCCAGTCCGATTTGCTTGGGGGGGCTGCTTTCTGTTCCGCTATGTACGTG	480		
Qy	439	-----TCCGAGCTGTTACCACTCTGCCTCGCCGCG	468		
Db	481	GGGGACCTCTCGGATCTGTCTTCCTCGTCTCCAGCTGTTTACCACATCTCGCCTCGCCGG	540		
Qy	469	CATGAGACGGTCAGGACGCAATTTGCTCAATCTATCCCGGCCACATAAACGGGTCAACCGT	528		
Db	541	CATGAGACGGTCAGGACGCAATTTGCTCAATCTATCCCGGCCACATAAACGGGTCAACCGT	600		
Qy	529	ATGCGTTGGGATATGATGATGAATCGTTCGCCCTACACAGCGCCCTGGTGGTATCGGAGCTG	589		
Db	601	ATGCGTTGGGATATGATGATGAATCGTTCGCCCTACACAGCGCCCTGGTGGTATCGGAGCTG	660		
Qy	589	CTCCGATCCCAACAGCTGTGTGGACATGTGGGGGGGGCCATTGGGGAGTCTCTGCG	648		
Db	661	CTCCGATCCCAACAGCTGTGTGGACATGTGGGGGGGGCCATTGGGGAGTCTCTGCG	720		
Qy	649	GGTCTCGCTACTATTCCATGGTGGGGAATCGGGCTTAAGTTTGTGATGCTACTC	708		
Db	721	GGTCTCGCTACTATTCCATGGTGGGGAATCGGGCTTAAGTTTGTGATGCTACTC	780		
Qy	709	TTTGCTCCCTTAATAG	723		
Db	781	TTTGCTCCCTTAATAG	795		
RESULT 8	AX685006	795 bp	DNA	linear	PAT 29-MAR-2003
LOCUS	Sequence 5 from Patent WO02055548.				
DEFINITION	AX685006				
ACCESSION	AX685006.1	GI:29371411			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					

```

/codon_start=1
/protein_id="CAD86520.1"
/db_xref="GI:29371412"
/translation="MLGKVIDTLGCFADLVGYIPLVGAFLGGAARALAHGVRVLEDG
VNYATNLPGCCFSFIPLALISCLITVPASAYEVNVMGVHVTIDNCNSI1VVEAADM
IMHTGCPVPCVRENNSRCRWALPTILAAKNASVPTTILRHVDLLGAAAFCSAMV
GLGCGSVFLVSGLFTISPRRHETVDCNCST1PGHITGHRMAMDMNNWSPPTALVVS
QLLRIPQAVVDVMVAGAHGWGLAGLAYISMVGNWAKVLLIVMLLFP"
1. .789
mat peptide

```



```
unclassified.
REFERENCE 1 (bases 1 to 2082)
AUTHORS Maertens,G., Bosman,F., De M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 47 15-FEB-1996;
COMMENT INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES
    source
        1..2082
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
        1..2079
            /note="unnamed protein product; Protein sequence is in conflict with the conceptual translation"
            /codon_start=1
            /protein_id="CAA03146.1"
            /db_xref="GI:2302423"
        /translation="NLGKIDVLTGCFADLVGYIPLVGPALGGAARALAHGVRLLEDG
        VNYATNLPFGCSFSLFLALLSCLTPASAYEVNRVSGMYHTVNDNSNSIYVEAADM
        IMHTPCGCVPCVRENNSSRCWALVPTLAARNASVPTTIRRHVDLLVGAARFCSAMYV
        GDLGSLVLSOLFSTISPRHETVODCNCSIYPGHLTGHMAWDMWMSPTTALVVS
        QLLRIPQAVDVAGAHGVLGALAYSNGVWAKVLVYMLLPAGVDGHTRVSGAAA
        SDRGLVSLFSPGSAQKIQLVNTGSHINRTALNCNDSLQTFGAALFYRHKFNSSG
        CPERLASCRSIDKFAQGGPLTYTFEENSQDRPYCWHYAPRCPGVPASQVCGPYCF
        TSPVVGTTDRFGVPTVMGANDSDVLIANNTRPRGNWFGCTWMNGTFTKTCGGP
        PCNIGAGNNTDTCTDCPRKPEATYARCGSGPWLTPRCMVHYVYRLWHYPTVNTF
        IPRVMYGVGVEHREACNWTGRCDLDRSELSPLLSLTTWEOILPCSPFTLP
        ALSTGLIHLQNVLYGVGSVAVSLVKEVYVLLFLLLADARACACLUWMLLI
        AQEAALENLVINAAVAGAHGTLSFLVFFCAAWYIKGRLVFGAAYFYGVWVPLLLI
        LLALPPRAYA"
        mat_peptide
            1..2076
                /product="unnamed"
                BASE COUNT 366 a 634 c 600 g 482 t
                ORIGIN
                    4 TTGGGTAAGGTTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63
                    DB 4 TTGGGTAAGGTTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63
                    64 CTCGTCGGCGCCCCCTAGCGGCGCTGCCAGGGCCCTGCGCATGCGCGGTTCG 123
                    DB 64 CTCGTCGGCGCCCCCTAGCGGCGCTGCCAGGGCCCTGCGCATGCGCGGTTCG 123
                    124 GAGGACGGCGTGAACCTATGACAGGAAATTTGCCCGGTTGCTTCTTCTATCTTCCTC 183
                    DB 124 GAGGACGGCGTGAACCTATGACAGGAAATTTGCCCGGTTGCTTCTTCTATCTTCCTC 183
                    184 TTGGCTTTGCTGTCGTCTGACAGTTCCTCAACTCAAGCAATTTGTCGCTTCTATCTTCCTC 243
                    DB 184 TTGGCTTTGCTGTCGTCTGACAGTTCCTCAACTCAAGCAATTTGTCGCTTCTATCTTCCTC 243
                    364 TTGGGTAGCGCTCAACCCCAAGCTCGAGCTAGGAAACCCAGCGTCCCAACCAAGCAATA 423
                    DB 364 TTGGGTAGCGCTCAACCCCAAGCTCGAGCTAGGAAACCCAGCGTCCCAACCAAGCAATA 423
                    424 CGACGCCACGTGCAT-----438
                    DB 424 CGACGCCACGTGCATTTGCTCGTGGGGCGGTGCTTTCTGTTCCGCTATGTACGTGGGG 483

Query Match 86.4%; Score 624.8; DB 6; Length 2082;
Best Local Similarity 90.5%; Pred. No. 2.1e-126;
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;

QY 4 TTGGGTAAGGTTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63
DB 4 TTGGGTAAGGTTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63
QY 64 CTCGTCGGCGCCCCCTAGCGGCGCTGCCAGGGCCCTGCGCATGCGCGGTTCG 123
DB 64 CTCGTCGGCGCCCCCTAGCGGCGCTGCCAGGGCCCTGCGCATGCGCGGTTCG 123
QY 124 GAGGACGGCGTGAACCTATGACAGGAAATTTGCCCGGTTGCTTCTTCTATCTTCCTC 183
DB 124 GAGGACGGCGTGAACCTATGACAGGAAATTTGCCCGGTTGCTTCTTCTATCTTCCTC 183
QY 184 TTGGCTTTGCTGTCGTCTGACAGTTCCTCAACTCAAGCAATTTGTCGCTTCTATCTTCCTC 243
DB 184 TTGGCTTTGCTGTCGTCTGACAGTTCCTCAACTCAAGCAATTTGTCGCTTCTATCTTCCTC 243
QY 364 TTGGGTAGCGCTCAACCCCAAGCTCGAGCTAGGAAACCCAGCGTCCCAACCAAGCAATA 423
DB 364 TTGGGTAGCGCTCAACCCCAAGCTCGAGCTAGGAAACCCAGCGTCCCAACCAAGCAATA 423
QY 424 CGACGCCACGTGCAT-----438
DB 424 CGACGCCACGTGCATTTGCTCGTGGGGCGGTGCTTTCTGTTCCGCTATGTACGTGGGG 483
```


304	ATGATCATGCACACCCCGGGTGGCTGCGTTCGGGAGAAACAACCTTCCCGCTGC	363
364	TGGTAGCGCTCACCCCAACCGCTCGCAGCTAGGAACGCCAGCGTCCCCACACGACAATA	423
364	TGGTAGCGCTCACCCCAACCGCTCGCAGCTAGGAACGCCAGCGTCCCCACACGACAATA	423
424	CGAGCGCCAGTTCGAT-----	438
424	CGAGCGCAAGTTCGATTTGCTCGTTGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGGG	483
439	-----TCCACAGCTGTTACACATCTCGCTCGCCGGGCAT	471
484	GACCTCTGGCGGATCTGTCTTCCTCGTCTCCAGCTGTTTACACATCTCGCTTCGCGGCAT	543
472	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAAACGGGTACCGGTATG	531
544	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAAACGGGTACCGGTATG	603
532	GCTTTGGGATATGATGAACCTGGTGCCTACAAACGGCCCTGGTGGTATCCAGACTGCTC	591
604	GCTTTGGGATATGATGAACCTGGTGCCTACAAACGGCCCTGGTGGTATCCAGACTGCTC	663
592	CGGATCCCAACAGCTGTCGTGGACATGTTGGCGGGGGCCCATTTGGGAGTCTCTGGCGGGT	651
664	CGGATCCCAACAGCTGTCGTGGACATGTTGGCGGGGGCCCATTTGGGAGTCTCTGGCGGGC	723
652	CTCGCTTACTATTCCATGTTGGGGAACCTGGGCTTAAGGTTTTGATTTGTGATGCTACTCTTT	711
724	CTCGCTTACTATTCCATGTTGGGGAACCTGGGCTTAAGGTTTTGATTTGTGATGCTACTCTTT	783
712	GC	713
784	GC	785

```

RESULT 11
AX452796
LOCUS      2082 bp      DNA
Sequence 47 from Patent EP1211315.
AX452796
AX452796.1 GI:21712481
Hepatitis C virus
Hepatitis C virus
Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.
Recombinant vectors for producing hcv envelope proteins
Patent: EP 1211315-A 47 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
Location/Qualifiers
1..2082
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
<1..2079
/feature="unnamed protein product"
/codon_start=1
/db_xref="GI:21712482"
/translation="NLGKVIDTLCGFDALVGVIPLVGAPLGGAARALAHGVRLVDG
VNYATGNLPGCSFSIFLLALLSCLTVPASAVEVNVSMGYHTVNDCSNSIIVCEAADM
IMHTPGVCPVRENSSRCWALPTPLAARNASVPTTIRHVDLLVGAAGAACSVAV
GDLCSGVLSQSFITISPRHETVQDCNSCIYSGHITGHRMANDMMNNSPTTALVVS
QLIRIPQAVDVMVAGAHWGLAGLAYSMVNGNKLAVMLLVFAGVDGHTRVSGGAA
SDTRGLVSLSPGSAOKIQLVNTNGSHWIIHRTALCNDLSLOTGFFAALFYKHFPNSG
CPERLACSRSIDFAQGWGPIITYEPNSDQRPYCMHYAPRPGCIVPASQGVGVTCG
TPSPVVGTTDRGVPYNNMGANDSVLILNNTPRPRGNFGCTWNGTGTFTCTCGP
PCNIGGAGNNLTCTCTDCFRKHPEATYARGSGPMLTPRCMVHPYRLMHPCTVNTPT
IFKRYMYGVGVEHFEACNWRGRCDEDRSELSPLILSTTQWIIILPCSTTULP
ALQGLTILHQNIVDVQYLVGVSALVSLIKWEIYVLLHLLFLLILLADARIACALNMWLLI
AQAFAALENLVVLNMAVAGAHGTSFLSVFCCAAMYIKGRLLVGAATYATGVQVPLLLI

```

[illegible]

RESULT 12					
AX685048					
LOCUS	AX685048	2082 bp	DNA	linear	PAT 29-MAR-2003
DEFINITION	Sequence 47 from Patent WO0205548.				
ACCESSION	AX685048				
VERSION	AX685048.1				
					GI:29371453

[illegible]

ORIGIN	Query Match Best Local Similarity Matches 708; Conservative	86.4%; Score 624.8; DB 6; Length 2433; 90.5%; Pred. No. 2.1e-126; 0; Mismatches - 2; Indels 72; Gaps 1;
QY	4	TTGGGTAAAGGTGATGATACCATCCCTTACATGCGGCTTTCGCCGACCTCGTGGGTGATACATCCG 63
Db	355	TTGGGTAAAGGTGATGATACCATCCCTTACATGCGGCTTTCGCCGACCTCGTGGGTGATACATCCG 414
QY	64	CTCGTCGGGCGCCCTTAGGGGGCGCTGTCACAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123
Db	415	CTCGTCGGGCGCCCTTAGGGGGCGCTGTCACAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 474
QY	124	GAGGACGGGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTCTC 183
Db	475	GAGGACGGGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTCTC 534
QY	184	TTGGCTTTTGCTGTCCTGTCTTGACCGTTTCCAGCTTCCGCTTATGAAGTGGCGAACGTGTCC 243
Db	535	TTGGCTTTTGCTGTCCTGTCTTGACCGTTTCCAGCTTCCGCTTATGAAGTGGCGAACGTGTCC 594
QY	244	GGGATGTACATGTACGACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCGGAC 303
Db	595	GGGATGTACATGTACGACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCGGAC 654
QY	304	ATGATCATGCACACCCCGGGTGGTGCCCTGGTTTCGGGAGAACCACTCTTCCCGCTGC 363
Db	655	ATGATCATGCACACCCCGGGTGGTGCCCTGGTTTCGGGAGAACCACTCTTCCCGCTGC 714
QY	364	TGGGTAGCGGTACCCCAAGCTCGAGCTAGGAGACGCGAGCTCCCAACGACGACATA 423
Db	715	TGGGTAGCGGTACCCCAAGCTCGAGCTAGGAGACGCGAGCTCCCAACGACGACATA 774
QY	424	CGAGCCACGTCGAT----- 438
Db	775	CGAGCCACGTCGATTTGCTCGTTGGGGGGGCTGCTTTTCTGTTCCGCTATGTACGTGGGG 834
QY	439	-----TCCGAGCTGTTCCACCTCTCGCTCGCTCCCGCGCAT 471
Db	835	GACCTCTCGGGATCTGTTCTTCTCGTCTCCAGCTGTTTCCACCTCTCGCTCTCCCGCGCAT 894
QY	472	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATPACGGGTCAACCGTATG 531
Db	895	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATPACGGGTCAACCGTATG 954
QY	532	GCTTGGGATATGATGATGAACTGTCGCTCAACAGGCGCTTGGGTATCGCAGCTGCTC 591
Db	955	GCTTGGGATATGATGATGAACTGTCGCTCAACAGGCGCTTGGGTATCGCAGCTGCTC 1014
QY	592	CGGATCCCAACAGCTGTCGTGACATGCTGGCGGGGCCATTTGGGAGTCTTGGCGGGT 651
Db	1015	CGGATCCCAACAGCTGTCGTGACATGCTGGCGGGGCCATTTGGGAGTCTTGGCGGGT 1074
QY	652	CTCGCTACTATTCATGATGTTGGGAACTGCGGCTAAGTTTGTATGTGATGCTACTCTTT 711
Db	1075	CTCGCTACTATTCATGATGTTGGGAACTGCGGCTAAGTTTGTATGTGATGCTACTCTTT 1134
QY	712	GC 713
Db	1135	GC 1136
RESULT 14		
AR157351	AR157351	2433 bp DNA linear PAT 17-OCT-2001
LOCUS	Sequence 49 from patent US 6245503.	
DEFINITION	AR157351	
ACCESSION	AR157351.1	
VERSION	GI:16218285	
KEYWORDS	Unknown.	
SOURCE	Unknown.	
ORGANISM	Unclassified.	

REFERENCE	1	(bases 1 to 2433)
AUTHORS	Maertens,G., Bosman,F.; De Martynoff,G. and Buyse,M.-A.	
TITLE	Purified Hepatitis C virus envelope proteins for diagnostic and therapeutic use	
JOURNAL	Patent: US 6245503-A 49 12-JUN-2001;	
FEATURES	Location/Qualifiers	
source	1..2433	
BASE COUNT	434 a 745 c 714 g 540 t	
ORIGIN	/organism="unknown"	
Query Match	86.4%; Score 624.8; DB 6; Length 2433;	
Best Local Similarity	90.5%; Pred. No. 2.1e-126;	
Matches 708; Conservative	0; Mismatches 2; Indels 72; Gaps 1;	
QY	4 TTGGTAAAGTCATGCATACCTTTACATCGCGCTTCGCCGACCTTCGTGGGGTACATTCCG 63	
Db	355 TTGGTAAAGTCATGCATACCTTTACATCGCGCTTCGCCGACCTTCGTGGGGTACATTCCG 414	
QY	64 CTGCTCGGCGCCCCCTAGGGGGCGCTGCAAGGGCCCTTGGGCGCATVGGCGTCCGGGTTCTG 123	
Db	415 CTGCTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTTGGGCGCATVGGCGTCCGGGTTCTG 474	
QY	124 GAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGTTGCTCTTCTCTATCTTCTC 183	
Db	475 GAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGTTGCTCTTCTCTATCTTCTC 534	
QY	184 TTGGCTTTGCTGTCCTGCTGACCGTTCCAGCTTCGGCTTATGAAAGTCGCAACGTGTCC 243	
Db	535 TTGGCTTTGCTGTCCTGCTGACCGTTCCAGCTTCGGCTTATGAAAGTCGCAACGTGTCC 594	
QY	244 GGGATGTACCATTGTCAAGAACGACTGCTCAAATCAAGCAATTTGTATGAGGACGGGAC 303	
Db	595 GGGATGTACCATTGTCAAGAACGACTGCTCAAATCAAGCAATTTGTATGAGGACGGGAC 654	
QY	304 ATGATCATGACACCCCCTGGTGGCTGCGTTCGGGGAACAACCTCTTCCCGCTGC 363	
Db	655 ATGATCATGACACCCCCTGGTGGCTGCGTTCGGGGAACAACCTCTTCCCGCTGC 714	
QY	364 TGGGTAGGCGTACACCCACGCTCGAGCTAGGAACGCGAGCTCCCAACACGACATA 423	
Db	715 TGGGTAGGCGTACACCCACGCTCGAGCTAGGAACGCGAGCTCCCAACACGACATA 774	
QY	424 CGACGCCACGTCGAT----- 438	
Db	775 CGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 834	
QY	439 -----TCCAGCTGTTTCAACATCTCGCTCGCCGGCAT 471	
Db	835 GACCTCTCGGATCTGTCTTCTCGTCTCCAGCTGTTTCAACATCTCGCTCGCCGGCAT 894	
QY	472 GAGAOGGTGACGAGCTGCAATTGCTCAATCTATCCCGGCCACATTAACGGGTACACGATG 531	
Db	895 GAGAOGGTGACGAGCTGCAATTGCTCAATCTATCCCGGCCACATTAACGGGTACACGATG 954	
QY	532 GCTTGGGATATGATGAACTGCTCGCTACACGGCCCTTGGTGTATCGCAGCTGCTC 591	
Db	955 GCTTGGGATATGATGAACTGCTCGCTACACGGCCCTTGGTGTATCGCAGCTGCTC 1014	
QY	592 CGGATCCCAACAGCTGCTGTGGACATGTTGGCGGGGCCCATTTGGGAGTCTTGGCGGGT 651	
Db	1015 CGGATCCCAACAGCTGCTGTGGACATGTTGGCGGGGCCCATTTGGGAGTCTTGGCGGGC 1074	
QY	652 CTCGCTACTATTCCATGTTGGGAACTAGGCTAAGTTTTGATTTGTGATGCTACTCTTT 711	
Db	1075 CTCGCTACTATTCCATGTTGGGAACTAGGCTAAGTTTTGATTTGTGATGCTACTCTTT 1134	
QY	712 GC 713	
Db	1135 GC 1136	
RESULT	15	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 199.88 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303a-21
Perfect score: 723
Sequence: 1 ATGTTGGTGAAGTCA...TACTCTTGTCTCCCTAATAG 723

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 19Jun03:*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	723	100.0	723	17	AAT12961 HCV E1 construct H
2	723	100.0	723	24	AAL48925 Hepatitis C virus
3	641	88.7	795	17	AAT12705 HCV E1 construct H
4	641	88.7	795	24	AAL48914 Hepatitis C virus
5	624.8	86.4	2082	24	AAL48939 Hepatitis C virus
6	624.8	86.4	2086	17	AAT12973 HCV E1 construct H
7	624.8	86.4	2433	17	AAT12974 HCV E1 construct H
8	613.8	84.9	2434	24	AAL48940 Hepatitis C virus

9	598.2	82.7	606	17	AAT12963 HCV E1 construct H
10	598.2	82.7	606	24	AAL48927 Hepatitis C virus
11	597	82.6	636	17	AAT12964 HCV E1 construct H
12	597	82.6	636	24	AAL48928 Hepatitis C virus
13	559.2	77.3	9605	24	ABK91411 Hepatitis C virus
14	559.2	77.3	9605	24	ABK91424 Hepatitis C virus
15	559.2	77.3	9605	24	ABK91425 Hepatitis C virus
16	559.2	77.3	9605	24	ABK91426 Hepatitis C virus
17	559.2	77.3	9605	24	ABK91428 Hepatitis C virus
18	559.2	77.3	9605	24	ABK91429 Hepatitis C virus
19	559.2	77.3	9605	24	ABK91430 Hepatitis C virus
20	559.2	77.3	9605	24	ABK91431 Hepatitis C virus
21	559.2	77.3	9605	24	ABK91432 Hepatitis C virus
22	559.2	77.3	9605	24	ABK91433 Hepatitis C virus
23	559.2	77.3	9605	24	AAD25332 Hepatitis C virus
24	559.2	77.3	9608	24	ABK91427 Hepatitis C virus
25	559.2	77.3	11062	24	AAD25331 Hepatitis C virus
26	559.2	77.3	11076	21	AAA98965 Hepatitis C virus
27	556	76.9	561	17	AAT12962 HCV E1 construct H
28	556	76.9	561	24	AAL48926 Hepatitis C virus
29	553.8	76.6	1251	13	AAQ26981 HCV gene 1. Hepat
30	553.8	76.6	3360	17	AAQ26981 Hepatitis C genome
31	553.8	76.6	3461	15	AAQ64068 Non-A, non-B hepat
32	553.8	76.6	3461	16	AAQ64068 S'UTR/CORE/ENV/NS1
33	553.8	76.6	9413	16	AAQ03960 Partial HCV non-st
34	553.8	76.6	9413	16	AAQ81559 Hepatitis C virus
35	553.8	76.6	9413	16	AAQ80498 DNA encoding HCV p
36	553.8	76.6	9413	24	AAD25517 Hepatitis C virus
37	553.8	76.6	9413	25	AAD49655 Hepatitis C virus
38	553.8	76.6	9413	25	AAL53723 Hepatitis C virus
39	552.8	76.5	1562	19	AAV60672 Fragment #5 isolat
40	552.8	76.5	1953	25	AAL55222 Plasmid pIDK2 DNA
41	552.8	76.5	2187	19	ABA03491 Cuticle protein 1
42	552.8	76.5	2540	14	AAQ43889 NAMH hepatitis vir
43	552.8	76.5	2540	15	AAQ63753 NAMHV genomic fra
44	552.8	76.5	2829	19	AAV60673 Fragment #6 isolat
45	552.2	76.4	9609	24	AAD33038 HCV-S1 full-length

ALIGNMENTS

RESULT 1
AAT12961
ID AAT12961 standard; DNA; 723 BP.

XX AAT12961;

XX 24-SEP-1996 (first entry)

XX HCV E1 construct HCCI37.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera;

XX Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP03031.

XX 29-JUL-1994; 94EP-0870132.

XX (INNO-) INNOGENETICS NV.

XX Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

61	CGCTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
121	CTGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTGTCTTTCTCTATCTTC	180
121	CTGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTGTCTTTCTCTATCTTC	180
181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTTCAGCTTCGGCTTATGAAGTGGCGAACGTG	240
181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTTCAGCTTCGGCTTATGAAGTGGCGAACGTG	240
241	TCCGGGATGTACCATGTACAGAAAGACTGTCCAACTCAAGCATTTGTATGAGCAGCG	300
241	TCCGGGATGTACCATGTACAGAAAGACTGTCCAACTCAAGCATTTGTATGAGCAGCG	300
301	GACATGATCATGCAACACCCCGGGTGGTGCCCTGGCTTCGGGAGAACAACTCTTCCCGC	360
301	GACATGATCATGCAACACCCCGGGTGGTGCCCTGGCTTCGGGAGAACAACTCTTCCCGC	360
361	TGCTGGGTAGCGTCAACCCACAGCTCGCAGTAGGAAGCCAGCGTCCCAACACAGACA	420
361	TGCTGGGTAGCGTCAACCCACAGCTCGCAGTAGGAAGCCAGCGTCCCAACACAGACA	420
421	ATACGACGCCACGTCGATTTCCACGCTGTTTCAACATCTCGCCCTCGCGGCGATGAGACGGTG	480
421	ATACGACGCCACGTCGATTTCCACGCTGTTTCAACATCTCGCCCTCGCGGCGATGAGACGGTG	480
481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCACCGTATGGCTGGGAT	540
481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCACCGTATGGCTGGGAT	540
541	ATGATGATGAACCTGGTGCCTTACAAACGGCCCTGTGGTATGCGACGTGTCGGATCCCA	600
541	ATGATGATGAACCTGGTGCCTTACAAACGGCCCTGTGGTATGCGACGTGTCGGATCCCA	600
601	CAAGCTGTGTGGACATGGTGGCGGGGGCCCATTTGGGGAGTCTCGCGGGTCTCGCCTAC	660
601	CAAGCTGTGTGGACATGGTGGCGGGGGCCCATTTGGGGAGTCTCGCGGGTCTCGCCTAC	660
661	TATTCCATGTGGGGAACTGGGCTAAAGTTTTTGATTGTGATGCTACTCTTTTGCTCCCTAA	720
661	TATTCCATGTGGGGAACTGGGCTAAAGTTTTTGATTGTGATGCTACTCTTTTGCTCCCTAA	720
721	TAG 723	
721	TAG 723	

RESULT 3

RESULTS	
3	
AAT12705	
XX	AAT12705 standard; DNA; 795 BP.
XX	
XX	AAT12705;
XX	
XX	AC
XX	AC
XX	23-SEP-1996 (first entry)
XX	
XX	HCV E1 construct HCC110A.
XX	
XX	HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX	serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX	ss.
XX	
XX	Hepatitis C virus.
OS	
XX	
XX	WO9604385-A2.
PN	
XX	
XX	15-FEB-1996.
PD	
XX	
XX	
XX	31-JUL-1995; 95WO-EP03031.
PF	
XX	
XX	29-JUL-1994; 94EP-0870132.
PR	
XX	
XX	(INNO-) INNOGENETICS NV.
PA	

QY	469	CATGAGCGGTGCAGGACTGCAATTGGCTCAATCTATCCCGGCCACATACGGGTCAACGGT	520
Db	541	CATGAGCGGTGCAGGACTGCAATTGGCTCAATCTATCCCGGCCACATACGGGTCAACGGT	600
QY	529	ATGGCTTGGGATATGATGATGAACCTGGTCGGCTACACGGCCCTGGTGGTATCGCAGCTG	588
Db	601	ATGGCTTGGGATATGATGATGAACCTGGTCGGCTACACGGCCCTGGTGGTATCGCAGCTG	660
QY	589	CTCCGGATCCCAACAGCTGTCTGTGGACATGCTGGCGGGGGCCCAATTGGGAGTCTTGGCG	648
Db	661	CTCCGGATCCCAACAGCTGTCTGTGGACATGCTGGCGGGGGCCCAATTGGGAGTCTTGGCG	720
QY	649	GGTCTCGCCTACTATTCCATGCTGGGAACTGGGCTAAGTTTTCATGTGATGCTACTTC	708
Db	721	GGTCTCGCCTACTATTCCATGCTGGGAACTGGGCTAAGTTTTCATGTGATGCTACTTC	780
QY	709	TTTGCTCCCTAATAG	723
Db	781	TTTGCTCCCTAATAG	795
RESULT 4			
AAL48914			
ID	AAL48914 standard; DNA; 795 BP.		
XX	AAL48914;		
XX	AC		
DT	24-OCT-2002 (first entry)		
XX	Hepatitis C virus clone HCC110A E1 protein coding sequence.		
DE	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;		
XX	virucide; immunostimulant; vaccine; ds.		
KW	Hepatitis C virus.		
XX	WO200255548-A2.		
XX	18-JUL-2002.		
XX	11-JAN-2002; 2002WO-EP00219.		
XX	11-JAN-2001; 2001US-260699P.		
PR	30-AUG-2001; 2001US-315768P.		
XX	(INNO-) INNOGENETICS NV.		
PA	Maertens G, Bosman F, Buysse M;		
XX	WPI; 2002-599657/64.		
XX	P-PSDB; AAO18661.		
XX	New therapeutic vaccine compositions comprising at least one purified		
PT	recombinant hepatitis C virus (HCV) single or specific oligomeric		
PT	recombinant envelope protein E1 or E2, useful for immunizing humans		
PT	from HCV infection		
XX	Example 2; Page 161-162; 243pp; English.		
PS	The present invention relates to new therapeutic vaccine compositions for		
XX	inducing hepatitis C virus (HCV)-specific antibodies, comprising a		
CC	composition containing at least one purified recombinant HCV single or		
CC	specific oligomeric recombinant envelope proteins selected from an E1 and		
CC	an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are		
CC	useful for inducing HCV-specific antibodies or for immunising humans		
CC	against HCV. The recombinant HCV E1 and/or E2 proteins are useful as		
CC	vaccines or therapeutics, in HCV screening and confirmatory antibody		
CC	tests, for raising antibodies, in the preparation of medicament, and for		
CC	in vitro monitoring of HCV disease or prognosing the response to		
CC	treatment of patients suffering from HCV infection. The present sequence		
CC	is a coding sequence described in the exemplification of the invention.		
XX	Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;		

KW virucide; immunostimulant; vaccine; ds.
 XX Hepatitis C virus.
 OS WO200255548-A2.
 PN 18-JUL-2002.
 XX 11-JAN-2002; 2002WO-BP00219.
 PF 11-JAN-2001; 2001US-260699P.
 XX 30-AUG-2001; 2001US-315768P.
 PR (INNO-) INNOGENETICS NV.
 PA Maertens G, Bosman F, Buyse M;
 XX WPI; 2002-599657/64.
 DR P-PSDB; AAO18678.
 XX New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans
 PT from HCV infection
 XX Example 2; Page 206-209; 243pp; English.
 PS The present invention relates to new therapeutic vaccine compositions for
 XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.
 XX Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 other;
 SQ

Query Match 86.4%; Score 624.8; DB 24; Length 2082;
 Best Local Similarity 90.5%; Pred. No. 1e-158;
 Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;

QY 4 TTGGGTAAGTTCATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGTACATTCCG 63
 Db 4 TTGGGTAAGTTCATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGTACATTCCG 63
 QY 64 CTCGTGGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGCGTCCGGGTTCTG 123
 Db 64 CTCGTGGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGCGTCCGGGTTCTG 123
 QY 124 GAGGACGGGTGAATATGCAACAGGAAATTTGCGCGGTGCTCTTTCTATCTTCCTC 183
 Db 124 GAGGACGGGTGAATATGCAACAGGAAATTTGCGCGGTGCTCTTTCTATCTTCCTC 183
 QY 184 TTGGCTTTCGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTTGCC 243
 Db 184 TTGGCTTTCGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTTGCC 243
 QY 244 GGGATGTACATGTGTCACAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
 Db 244 GGGATGTACATGTGTCACAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
 QY 304 ATGATCATGCAACACCCCGGGTGGTGGCTTCGCTTGGGAGAACAACTCTTCCCGCTGC 363
 Db 304 ATGATCATGCAACACCCCGGGTGGTGGCTTCGCTTGGGAGAACAACTCTTCCCGCTGC 363
 QY 364 TGGGTAGCGCTCACCCCGGCTCGCGCTAGGACCGGCTCCCGACCAATA 423
 Db 364 TGGGTAGCGCTCACCCCGGCTCGCGCTAGGACCGGCTCCCGACCAATA 423

QY 424 CGACGCCACGTCGAT----- 438
 Db 424 CGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
 QY 439 -----TCCCAAGCTGTTCCACCATCTCGCCTCGCGGCAT 471
 Db 484 GACCTCTCGGGATCTGCTTCTCGTCTCCAGCTGTTCCACCATCTCGCCTCGCGGCAT 543
 QY 472 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTACCGTATG 531
 Db 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTACCGTATG 603
 QY 532 GCTTGGGATATGATGATGAACTGCTGCGCTCAACAGGCCCTGGTGTATCGCAGTGTCTC 591
 Db 604 GCTTGGGATATGATGATGAACTGCTGCGCTCAACAGGCCCTGGTGTATCGCAGTGTCTC 663
 QY 592 CGGATCCCAACAGCTGCTGTCGACATGTTGGGGGGGCCCATTTGGGAGTCTCTGGCGGT 651
 Db 664 CGGATCCCAACAGCTGCTGTCGACATGTTGGGGGGGCCCATTTGGGAGTCTCTGGCGGC 723
 QY 652 CTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 711
 Db 724 CTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 783
 QY 712 GC 713
 Db 784 GC 785

RESULT 6
 AAT12973
 ID AAT12973 standard; DNA; 2086 BP.
 XX AAT12973;
 DT 24-SEP-1996 (first entry)
 XX HCV E1 construct HCCI65.
 XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 XX OS Hepatitis C virus.
 PN WO9604385-A2.
 XX PD 15-FEB-1996.
 PF 31-JUL-1995; 95WO-BP03031.
 XX PR 29-JUL-1994; 94EP-0870132.
 XX (INNO-) INNOGENETICS NV.
 XX Bosman F, Buyse M, De Martynoff G, Maertens G;
 XX WPI; 1996-129401/13.
 DR Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 PT proteins - in presence of disulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX Claim 23; Fig 21; 146pp; English.
 XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a

disulphide bond cleavage agent, after lysis of recombinant host cells.
The constructs containing the purified HCV envelope proteins can be used
for vaccinating humans against HCV, for in vitro detection of HCV
antibodies in a sample, and in a serotyping assay for detecting one or
more serological types of HCV present in a biological sample. The
constructs can also be immobilised on a solid substrate and incorporated
into a reversed phase hybridisation assay for determining the presence or
the genotype of HCV. The new purification method preserves the
conformation of the recombinantly expressed E1, E2 and E1/E2, and
eliminates contaminating proteins. Antigens isolated using this method
are more reactive with human sera than those isolated by known
techniques.

XX Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;
SQ

Query Match 86.4%; Score 624.8; DB 17; Length 2086;
Best Local Similarity 90.5%; Pred. No. 1e-158; 2; Indels 72; Gaps 1;
Matches 708; Conservative 0; Mismatches 2;

4 TTGGTAAAGGTATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGGTACATTCCG 63
4 TTGGTAAAGGTATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGGTACATTCCG 63
64 CTCGTGGCGCCCTAGGGGCGCTGCGAGGCGCTGGCGATGGCGTCCGGGTTCTG 123
64 CTCGTGGCGCCCTAGGGGCGCTGCGAGGCGCTGGCGATGGCGTCCGGGTTCTG 123
124 GAGACGGGTGAACTATGCAACAGGGAATTCGCCGGTTGCTTCTATCTTCCTC 183
124 GAGACGGGTGAACTATGCAACAGGGAATTCGCCGGTTGCTTCTATCTTCCTC 183
184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTGTC 243
184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTGTC 243
244 GGGATGACCATGTCAGCAAGAGTCTCACTCAAGCAATTCGTATGAGGCGAGCGAC 303
244 GGGATGACCATGTCAGCAAGAGTCTCACTCAAGCAATTCGTATGAGGCGAGCGAC 303
304 ATGATATGACACACCCCGGGTGGTCCCTGCTCGGAGAACAACTTCCCGCTGC 363
304 ATGATATGACACACCCCGGGTGGTCCCTGCTCGGAGAACAACTTCCCGCTGC 363
364 TTGGTAGCGCTCACCCACAGCTCGAGTAGGAACCGCGTCCACACAGCAATA 423
364 TTGGTAGCGCTCACCCACAGCTCGAGTAGGAACCGCGTCCACACAGCAATA 423
424 CGAGCCACGTCGAT----- 438
424 CGAGCCACGTCGAT----- 438
439 -----TCCAGGTGTTTACCATCTGCTCGCTCGCGGAT 471
484 GACCTCTGGGATCTGCTTCTGCTCTCCAGCTGTTTCCACATCTCGCTCGCGGAT 543
472 GAGACGGTGCAGACTGCAATGCTCAATCTATCCCGCCACATACCGGTTCACGATG 531
544 GAGACGGTGCAGACTGCAATGCTCAATCTATCCCGCCACATACCGGTTCACGATG 603
532 GCTTGGGATATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 591
604 GCTTGGGATATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
592 CGGATCCCAAGCTCTGTCGACATGTTGGCGGGGCCCATTTGGGGAGTCTTCGGGGT 651
664 CGGATCCCAAGCTCTGTCGACATGTTGGCGGGGCCCATTTGGGGAGTCTTCGGGGG 723
652 CTCGCTACTATTCCATGTTGGGAACTGCGGCTTAAGGTTTGTGATGCTACTCTTT 711
724 CTCGCTACTATTCCATGTTGGGAACTGCGGCTTAAGGTTTGTGATGCTACTCTTT 783
712 GC 713
||

784 GC 785

RESULT 7
AAT12974
ID AAT12974 standard; DNA; 2433 BP.
XX
AC AAT12974;
XX
DT 25-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC166.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
FN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
XX WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di-sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by a
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.

XX Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;
SQ

Query Match 86.4%; Score 624.8; DB 17; Length 2433;
Best Local Similarity 90.5%; Pred. No. 1.1e-158;
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;

4 TTGGTAAAGGTATCGATACCCCTTACATGGGCTTCGCGACCTCGTGGGGTACATTCCG 63
355 TTGGTAAAGGTATCGATACCCCTTACATGGGCTTCGCGACCTCGTGGGGTACATTCCG 414
64 CTCGTGGCGCCCTAGGGGCGCTGCGAGGCGCTGCGAGGCGCTGCGAGGCGCTGCGGTTCTG 123
415 CTCGTGGCGCCCTAGGGGCGCTGCGAGGCGCTGCGAGGCGCTGCGGTTCTG 474

531 GGCTTTGGGATATGATGAACTGGTCCCTACACGGCCCTGGTATCGCAGCTGCT 590
1014
955 GGCTTTGGGATATGATGAACTGGTCCCTACACGGCCCTGGTATCGCAGCTGCT 1014
591 CGGGATCCCAAGCTGTCGTGGACATGCTGGCGGGGCCCATTTGGGGAGTCCTTGGCGGG 650
1015 CGGGATCCCAAGCTGTCGTGGACATGCTGGCGGGGCCCATTTGGGGAGTCCTTGGCGGG 1074
651 TCTCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTGTATGTGATGCTACTCTT 710
1075 CCTCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTGTATGTGATGCTACTCTT 1134
711 TGC 713
1135 TGC 1137

RESULT 9
AAT12963
ID AAT12963 standard; DNA; 606 BP.
XX
AC AAT12963;
DT 24-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC139.
XX
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
XX WO9604385-A2.
XX
XX 15-FEB-1996.
XX
XX 31-JUL-1995; 95WO-EP03031.
XX
XX 29-JUL-1994; 94EP-0870132.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
XX WPI; 1996-129401/13.

Claim 23; Fig 21; 146pp; English.
XX
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or
XX the genotype of HCV. The new purification method preserves the
XX conformation of the recombinantly expressed E1, E2 and E1/E2, and
XX eliminates contaminating proteins. Antigens isolated using this method
XX are more reactive with human sera than those isolated by known
XX techniques.

SQ Sequence 606 BP; 109 A; 193 C; 167 G; 137 T; 0 other;
Query Match 82.7%; Score 598.2; DB 17; Length 606;
Best Local Similarity 99.5%; Pred. No. 1.1e-151;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAAGTCAATGATACCTTACATCGCGCTTCGCGGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCAATGATACCTTACATCGCGCTTCGCGGACCTCGTGGGGTACATT 60
QY 61 CCCTCGTGGGGCCCCCCCCCTAGGGGCGCTGCGAGGCGCTCGCGGATGCGCTCGGGTT 120
DB 61 CCCTCGTGGGGCCCCCCCCCTAGGGGCGCTGCGAGGCGCTCGCGGATGCGCTCGGGTT 120
QY 121 CTGGAGGACGGCGTGAATCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAATCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGCTCTGACCGTTTCCAGCTTCCGCTTATGAAGTGGCAACGTTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGCTCTGACCGTTTCCAGCTTCCGCTTATGAAGTGGCAACGTTG 240
QY 241 TCCGGGATGTACCATGTGACGAAACGACTGCTCAACTCAAGCATTTGTATGAGGACGCG 300
DB 241 TCCGGGATGTACCATGTGACGAAACGACTGCTCAACTCAAGCATTTGTATGAGGACGCG 300
QY 301 GACATGATCATGACACCCCGGGTGGTGGCTTCCGCTTCCGCTTATGAAGTGGCAACGTTG 360
DB 301 GACATGATCATGACACCCCGGGTGGTGGCTTCCGCTTCCGCTTATGAAGTGGCAACGTTG 360
QY 361 TGCTGGGTAGCGCTCACCCCGACGCTCGAGCTAGGAAACGCGGCTCCCGGATGAGACG 420
DB 361 TGCTGGGTAGCGCTCACCCCGACGCTCGAGCTAGGAAACGCGGCTCCCGGATGAGACG 420
QY 421 ATAGAGCCGACGCTCGATTTCCAGCTGTTCAACATCTCGCTCGCGGATGAGACG 480
DB 421 ATAGAGCCGACGCTCGATTTCCAGCTGTTCAACATCTCGCTCGCGGATGAGACG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCATTAACGGGTCAACGGTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCATTAACGGGTCAACGGTTGGGAT 540
QY 541 ATGATGATGAATGCTGCTGCTTACACGCGCTTACAGCGGCTTGGTATCGCAGCTGCTCCGGATCCCA 600
DB 541 ATGATGATGAATGCTGCTGCTTACACGCGCTTACAGCGGCTTGGTATCGCAGCTGCTCCGGATCCCA 600
QY 601 CAA 603
DB 601 TAA 603

RESULT 10
AAL48927
ID AAL48927 standard; DNA; 606 BP.
XX
AC AAL48927;
DT 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus clone HCC139 E1 protein coding sequence.
XX
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.
XX
XX Hepatitis C virus.
XX
XX WO200255548-A2.
XX
XX 18-JUL-2002.
XX
XX 11-JAN-2002; 2002WO-EP00219.
XX
XX 11-JAN-2001; 2001US-260699P.
PR


```
PR 30-AUG-2001; 2001US-315768P.
XX (INNO-) INNOGENETICS NV.
XX Maertens G, Bosman F, Buyse M;
XX WPI: 2002-599657/64.
XX P-PSDB; AAO18668.
XX New therapeutic vaccine compositions comprising at least one purified
XX recombinant hepatitis C virus (HCV) single or specific oligomeric
XX recombinant envelope protein E1 or E2, useful for immunizing humans
XX from HCV infection
XX
XX Example 2; Page 177-178; 243pp; English.
XX
XX The present invention relates to new therapeutic vaccine compositions for
XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
XX composition containing at least one purified recombinant HCV single or
XX specific oligomeric recombinant envelope proteins selected from an E1 and
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
XX useful for inducing HCV-specific antibodies or for immunising humans
XX against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
XX vaccines or therapeutics, in HCV screening and confirmatory antibody
XX tests, for raising antibodies, in the preparation of medicament, and for
XX in vitro monitoring of HCV disease or prognosing the response to
XX treatment of patients suffering from HCV infection. The present sequence
XX is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 606 BP; 109 A; 193 C; 167 G; 137 T; 0 other;
XX
XX Query Match 82.7%; Score 598.2; DB 24; Length 606;
XX Best Local Similarity 99.5%; Pred. No. 1.1e-151;
XX Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 ATGTTGGGTAAGGTCAATCATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB |||||
DB 1 ATGTTGGGTAAGGTCAATCATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CGCTCTGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGCGGTGTT 120
DB |||||
DB 61 CGCTCTGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGCGGTGTT 120
QY 121 CTGGAGGACGGGTGAATCATGCAACAGGGAATTTGCCCGGTTCCTTTCTATCTTC 180
DB |||||
DB 121 CTGGAGGACGGGTGAATCATGCAACAGGGAATTTGCCCGGTTCCTTTCTATCTTC 180
QY 181 CTCCTGGCTTGTGTCCTGTGACCGCTTCAGCTTCGGCTTATGAAGTGGCGCAACGTG 240
DB |||||
DB 181 CTCCTGGCTTGTGTCCTGTGACCGCTTCAGCTTCGGCTTATGAAGTGGCGCAACGTG 240
QY 241 TCCGGATGTACCATGTCAAGACGACTGCTCAACTCAAGCATTCGTATGAGGCGCG 300
DB |||||
DB 241 TCCGGATGTACCATGTCAAGACGACTGCTCAACTCAAGCATTCGTATGAGGCGCG 300
QY 301 GACATGATCATGACACCCCGGGTCCGTCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
DB |||||
DB 301 GACATGATCATGACACCCCGGGTCCGTCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGGTCAACCCCGAGCTTCGAGCTTAGGAACCCAGCGTCCCGACGACA 420
DB |||||
DB 361 TGTGGGTAGCGGTCAACCCCGAGCTTCGAGCTTAGGAACCCAGCGTCCCGACGACA 420
QY 421 ATACGAGCGCATGTCATTCACGCTTCACCATCTCGCTCGCGGATGAGCGGTG 480
DB |||||
DB 421 ATACGAGCGCATGTCATTCACGCTTCACCATCTCGCTCGCGGATGAGCGGTG 480
QY 481 CAGGACTGCAATGCTCAATCTATCCCGGCCACATAACGGGTCAACGTTGGCTTGGAT 540
DB |||||
DB 481 CAGGACTGCAATGCTCAATCTATCCCGGCCACATAACGGGTCAACGTTGGCTTGGAT 540
QY 541 ATGATGATGAATGCTCGCTTACACGGCCCTGGTGTATCGCAGCTGCTCGGATCCCA 600
DB |||||
DB 1 ATGTTGGGTAAGGTCAATCATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAGGTCAATCATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

Db 541 ATGATGATGAATGCTCGCTTACACGGCCCTGGTGTATCGCAGCTGCTCGGATCCCTC 600
QY 601 CAA 603
Db 601 TAA 603

RESULT 11
AAT12964
ID AAT12964 standard; DNA; 636 BP.
XX AAT12964;
XX 24-SEP-1996 (first entry)
XX HCV E1 construct HCC140.
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX Hepatitis C virus.
XX WO9604385-A2.
XX 15-FEB-1996.
XX 31-JUL-1995; 95WO-EP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
XX proteins - in presence of disulphide bond cleavage agent, to
XX produce proteins suitable for direct use in vaccines or diagnostic
XX assays of HCV
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or
XX the genotype of HCV. The new purification method preserves the
XX conformation of the recombinantly expressed E1, E2 and E1/E2, and
XX eliminates contaminating proteins. Antigens isolated using this method
XX are more reactive with human sera than those isolated by known
XX techniques.
XX Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;
XX
XX Query Match 82.6%; Score 597; DB 17; Length 636;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-151;
XX Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ATGTTGGGTAAGGTCAATCATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB |||||
DB 1 ATGTTGGGTAAGGTCAATCATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
```


QY	61	CGGCTCGTCGGCGCCCCCTAGGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
Db	61		
QY	121	CTGGAGGACGGGTGAACATATGCAACAGGAAATTGCCGGTTGCTTCTTCTCTATCTTC	180
Db	121		
QY	181	CTCTGGCTTTGCTGCTCTGCTGTGACGGTTCCAGTTCGCGTTTCAAGTGGCGCAACGTTG	240
Db	181		
QY	241	TCGGGATGTACATGTACGAAACGACTGCTCCAACTCAAGCAATTGTGTATGAGCGAGCG	300
Db	241		
QY	301	GACATCATATGCACACCCCGGGTGGCGTCCCTGCGTTTGGGAGAACAACTTCTCCCGCG	360
Db	301		
QY	361	TGCTGGGTAGCGTCAACCCCAAGCTCGGAGTAGGAACGCCAGCGTCCCGCACAGACA	420
Db	361		
QY	421	ATACGAGCGACGTCGATTTCCAGCTGTTTACCATCTCGCCTCGCCGGCATGAGACGGTG	480
Db	421		
QY	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTTGGGAT	540
Db	481		
QY	541	ATGATGATGAATCGGTGCGCTACAAACGGCGCTGTGGTATCGCAGCTGCTCCGGATC	597
Db	541		

RESULT 12	
AAI48928	
ID	AAI48928 standard; DNA; 636 BP.
XX	
AC	AAI48928;
XX	
DT	24-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus clone HCCI40 E1 protein coding sequence.
XX	
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW	virucide; immunostimulant; vaccine; ds.
XX	
OS	Hepatitis C virus.
XX	
PN	WO200255548-A2.
XX	
PD	18-JUL-2002.
XX	
PF	11-JAN-2002; 2002WO-EP00219.
XX	
PR	11-JAN-2001; 2001US-260699P.
PR	30-AUG-2001; 2001US-315768P.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Maertens G, Boeman F, Buyse M;
XX	
DR	WPI: 2002-599657/64.
DR	P-PSDB; AAO18669.
XX	
PT	New therapeutic vaccine compositions comprising at least one purified
PT	recombinant hepatitis C virus (HCV) single or specific oligomeric
PT	recombinant envelope protein E1 or E2, useful for immunizing humans
PT	from HCV infection -
XX	
XX	

Example 2; Page 179-180; 243pp; English.

The present invention relates to new therapeutic vaccine compositions for inducing hepatitis C virus (HCV)-specific antibodies, comprising a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope proteins selected from an E1 and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are useful for inducing HCV-specific antibodies or for immunising humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, for raising antibodies, in the preparation of medicament, and for in vitro monitoring of HCV disease or prognosing the response to treatment of patients suffering from HCV infection. The present sequence is a coding sequence described in the exemplification of the invention.

Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;

Query Match 82.6%; Score 597; DB 24; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.3e-151;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGGGTAAGTCATCGATACCCCTACATCGGCTTCCGCCACTCGTGCGGTACATT 60
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 ATGTGGGTAAGTCATCGATACCCCTACATCGGCTTCCGCCACTCGTGCGGTACATT 60

QY 61 CGGCTGTCGCGCGCCCCCTTAGGGGGCGCTGCAGGGCCCTCGCGCATGGGCTCCGGTT 120
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 CGGCTGTCGCGCGCCCCCTTAGGGGGCGCTGCAGGGCCCTCGCGCATGGGCTCCGGTT 120

QY 121 CTGAGGACGCGCTGAATATGCAAAGGAAATTTCCCGGTTGCTTTCTCTATCTTC 180
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 CTGAGGACGCGCTGAATATGCAAAGGAAATTTCCCGGTTGCTTTCTCTATCTTC 180

QY 181 CTCCTTGGCTTGGCTGCTCTGCTCAGCTTTCAGCTTTCGCTTATGAAGTCGCAACG 240
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
181 CTCCTTGGCTTGGCTGCTCTGCTCAGCTTTCAGCTTTCGCTTATGAAGTCGCAACG 240

QY 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300

QY 301 GACATGATCATGCACACCCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 GACATGATCATGCACACCCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360

QY 361 TGCTGGGTAGCGCTCACCCCCAGCTCGAGCTAGGAAACGCGAGCTCCACCACGACA 420
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
361 TGCTGGGTAGCGCTCACCCCCAGCTCGAGCTAGGAAACGCGAGCTCCACCACGACA 420

QY 421 ATACGACGCGACGTCGATTCACGCTGTTCCACCATCTCGCTCGCGGGCATGAGACGG 480
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
421 ATACGACGCGACGTCGATTCACGCTGTTCCACCATCTCGCTCGCGGGCATGAGACGG 480

QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGG 540
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGG 540

QY 541 ATGATGATGAACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 597
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 ATGATGATGAACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 597

RESULT 13
ABK91411
ID ABK91411 standard; DNA; 9605 BP.
XX ABK91411;
AC ABK91411;
DX 15-NOV-2002 (first entry)
XX Hepatitis C virus Con 1 isolate DNA.
DE HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;

816 GAGGACGGC

816 GAGGACGGCGTGAACTATGCAACAGGGAATCTGCCCGGTGCTCCTTTTCTATCTTCCTT 875

PN WO200259321-A2.

01-AUG-2002.

16-JAN-2002; 2002WO-EP00526.

23-JAN-2001; 2001US-263479P.

(RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.

De Francesco R, Migliaccio G, Paonessa G;

WPI; 2002-599793/64.

New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region, useful in studying HCV replication and expression -

Claim 9; Page -: 69pp; English.

The invention relates to nucleic acid molecules comprising altered HCV NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region coding for one or more NS3, NS5A, or EMCV IRES mutations, respectively. The location of the mutations are detailed in the specification. Also included are (1) an expression vector comprising a nucleotide sequence coding for the altered nucleic acids, which is transcriptionally coupled to an exogenous promoter; (2) a recombinant cell human hepatoma cell comprising the altered nucleic acids; (3) a recombinant cell produced by introducing into a human hepatoma cell the altered nucleic acids; (4) producing a functional HCV replicon; (5) an HCV replicon enhanced cells made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicons and HCV replicon enhanced cells are useful in studying HCV replication and expression, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated diseases such as liver failure, cirrhosis and hepatocellular carcinoma. The present sequence is an HCV replicon Con 1 mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the HCV sequence appearing as ABK91411 and the information in Claim 9.

Sequence 9605 BP; 1910 A; 2884 C; 2732 G; 2079 T; 0 other;

Query Match 77.3%; Score 559.2; DB 24; Length 9605;
Best Local Similarity 85.3%; Pred. No. 8.9e-141;
Matches 667; Conservative 0; Mismatches 43; Indels 72; Gaps 1;

QY 4 TTGGGTAGGTCATCGATACCTTACATCGGGCTTCGCCGACTCGTGCGGTACATTCCG 63
Db 696 TTGGGTAGGTCATCGATACCTTACATCGGGCTTCGCCGACTCATGGGTCATTCGG 755

QY 64 CTCGTGGCGCCCCCTCAGGGGCGCTCCAGGGCCCTGGCGATGCCTCCGGTTCG 123
Db 756 CTCGTGGCGCCCCCTCAGGGGCGCTCCAGGGCCCTGGCGATGCCTCCGGTTCG 815

QY 124 GAGGACGGCGTGAACCTATGCAACAGGAAATTTGCCCGGTTCCTCTTTCTATTCCTC 183
Db 816 GAGGACGGCGTGAACCTATGCAACAGGAAATTCGCCCGGTTCCTCTTTCTATTCCTC 875

QY 184 TTGGCTTTGCTGTCCTGTCTGACCGTTTCAGCTTCGGCTTATGAAGTCGCAACGTCTC 243
Db 876 TTGGCTTTGCTGTCCTGTGTAACCATCCAGCTTCGGTTCATGAAGTCGCAACGTATCC 935

QY 244 GGAGTGTACATGTCACGAACGACTGCTCCAACCTCAAGCAATTTGTATGAGCGACGGAC 303
Db 936 GGAGTGTACATGTCACGAACGACTGCTCCAACGCAAGCAATTTGTATGAGCGACGGAC 995

QY 304 ATGATCATGCACACCCCGGGTCGTGCCCTCGGTTCGGGAGAACCACTCTTCCCGTGC 363
Db 996 ATGATCATGCATACCCCGGGTCGTGCCCTCGGTTCGGGAGAACCACTCTTCCCGTGC 1055

XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression
XX
PS Claim 9; Page -: 69pp; English.
XX
CC The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is an HCV replicon Con 1 mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the HCV sequence appearing as
CC ABK91411 and the information in Claim 9.
XX
SQ Sequence 9605 BP; 1909 A; 2883 C; 2734 G; 2079 T; 0 other;

Query Match 77.3%; Score 559.2; DB 24; Length 9605;
Best Local Similarity 85.3%; Pred. No. 8.9e-141;
Matches 667; Conservative 0; Mismatches 43; Indels 72; Gaps 1;

QY 4 TTGGGTAAAGTTCATCGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 63
DB 696 TTGGGTAAAGTTCATCGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 755
QY 64 CTCGTGGCGCCCGCTAGGGGGCGGTGCGAGGGCCCTGGCGATGGCTCGGGTCTG 123
DB 756 CTCGTGGCGCCCGCTAGGGGGCGGTGCGAGGGCCCTGGCGATGGCTCGGGTCTG 815
QY 124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTCCTC 183
DB 816 GAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTCCTC 875
QY 184 TTGGCTTTTGTCTGCTGCTGACCGGTTCAGCTTCGCGTTATGAAGTGGCAACGTGTC 243
DB 876 TTGGCTTTTGTCTGCTGCTGACCGGTTCAGCTTCGCGTTATGAAGTGGCAACGTATCC 935
QY 244 GGGATGTACATGTACGAAAGCTGCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 303
DB 936 GGAGTGTACATGTACGAAAGCTGCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 995
QY 304 ATGATCATGCACACCCCGGTGGTGGCTTGGTGGGAGAACACTTCTCCCGCTGC 363
DB 996 ATGATCATGCATACCCCGGTGGTGGCTTGGTGGGAGAACACTTCTCCCGCTGC 1055
QY 364 TGGGTAGCGCTCAACCCCGGTGGTGGCTTGGTGGGAGAACACTTCTCCCGCTGC 423
DB 1056 TGGGTAGCGCTCACTCCAGCGTTCGCGGCGAGAACCTAGCGTCCCACTACGACGATA 1115
QY 424 CGACGCCACGTGAT- 439
DB 1116 CGACGCCACGTGATTTGCTGTTGGGGGGGTGCTCTCTGCTCCGCTATGTACGTGGGA 1175
QY 440 -----CCAGCTGTTTACCATCTCGCTCGCGGCAAT 471
DB 1176 GATCTCTGCGGATCTGTTTTCCTGCTGCGGCCAGCTGTTTACCTTCTCGCTCGCGGCAAT 1235

Search completed: December 19, 2003, 18:51:12
Job time : 203.88 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1859.96 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-21
Perfect score: 723
Sequence: 1 ATGTTGGTAAAGTCATCGA.....TACTCTTCTCCCTAATAG 723

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estnu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_hic:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_hic:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pin:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_pbg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	101.2	14.0	488	9	AV755731
C 2	79.6	11.0	492	9	AV758366
C 3	41.6	5.8	502	12	BI879124
C 4	40.6	5.6	275	9	AV835132

5	40.6	5.6	402	9	AV392783
6	40.6	5.6	551	9	AV392165
7	40.6	5.6	552	12	BI996341
8	40.6	5.6	584	12	BI727879
9	40.4	5.6	1201	13	BX356664
10	40.2	5.6	1162	12	BM918259
11	40.2	5.5	1201	9	AL513886
C 11	39	5.4	359	12	BJ252669
C 12	39	5.4	375	12	BJ246716
C 13	39	5.4	840	29	CC335916
C 14	39	5.4	873	14	CD446071
C 15	39	5.4	873	14	CD446071
C 16	38.6	5.3	925	29	CNS0091P
C 17	38.4	5.3	636	12	BI960110
C 18	38.4	5.3	702	14	CD432549
C 19	38.4	5.3	740	12	BJ536071
C 20	38.4	5.3	925	29	CNS0091P
C 21	38.4	5.3	970	29	CNS010C9
C 22	38.4	5.3	987	29	CNS015VX
C 23	38.2	5.3	533	6	AU192776
C 24	38.2	5.3	538	6	AU193705
C 25	38.2	5.3	544	6	AU190971
C 26	38.2	5.3	544	6	AU192419
C 27	38.2	5.3	1270	12	BG968359
C 28	38	5.3	354	14	CB966525
C 29	38	5.3	742	13	BQ752673
C 30	38	5.3	1201	13	BX381961
C 31	37.8	5.2	435	14	C72860
C 32	37.8	5.2	533	29	CC010084
C 33	37.8	5.2	659	29	CC405164
C 34	37.8	5.2	826	29	BZ736582
C 35	37.8	5.2	895	29	CC359028
C 36	37.8	5.2	925	29	CC359026
C 37	37.8	5.2	940	29	CC010085
C 38	37.8	5.2	951	29	CC405167
C 39	37.6	5.2	431	9	AV639153
C 40	37.4	5.2	360	9	AJ473805
C 41	37.4	5.2	637	13	BQ293470
C 42	37.4	5.2	641	13	BQ172543
C 43	37.4	5.2	650	14	CA828039
C 44	37.4	5.2	834	29	BZ641450
C 45	37.4	5.2	841	29	BZ641457

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION AV755731 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
TITLE Homo sapiens cDNA BM clones
JOURNAL Unpublished
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES


```

source
1..488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sfIIA; Site_2: sfIIIB"
BASE COUNT 116 a 134 c 137 g 97 t
ORIGIN
Query Match 14.0%; Score 101.2; DB 9; Length 488;
Best Local Similarity 65.9%; Pred. No. 2.7e-15; Indels 1; Gaps 1;
Matches 162; Conservative 0; Mismatches 83;
445 CTGTTACCACTCTCGCTCGCGCATGAGACGGTGCAGGACTGCAATTTGCTCAATCTAT 504
403 CAGCTGATCATCTGGCTCAGCACCATGAGTTTGTGCATGAATGCAACTGCTCCATCTAT 344
505 CCCGGCCACATAACGGGTACCGTATG-GCTTGGGATATGATGAACTGCTGCGCTAC 563
343 CTTGGGCCATCATCTGACACCGTATGAGCATGGGACATGATGAATGCTGTGTCAC 284
564 AACGGCCCTCGTGTATCGAGCTGCTCGGATCCCAAGCTGCTGCGACATGCTGGC 623
283 CGCTGCTATGATCATGCGGTACGCAATGCGGTTCTGAGGTCTCATATAGATATCATCAG 224
624 GGGGGCCCATGCGGAGTCTCGCGGCTCTCGCTACTATTCATTCATGCGGAACTGGGC 683
223 CGGGGCTCACTGGGCGCTCATGTTTCGGCTTAGCTACTTCTCTATGACGAGCGTGGGC 164
684 TAAGCT 689
163 GAAAGT 158

RESULT 2
AV758366/c
LOCUS AV758366 BM Homo sapiens cDNA clone BMPAKA03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 492)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1..492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAKA03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"

/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sfIIA; Site_2: sfIIIB"
BASE COUNT 124 a 128 c 125 g 112 t
ORIGIN
Query Match 11.0%; Score 79.6; DB 9; Length 492;
Best Local Similarity 60.7%; Pred. No. 8.8e-10;
Matches 147; Conservative 0; Mismatches 94; Indels 1; Gaps 1;
449 TCACCATCTCGCTCGCGCATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCG 508
400 TGATTATCTCTCAGCAGCAACATTGGTTTGTGCAAGATGCAACTGCTCATTTCTATCCTG 341
509 GCCACATAACGGGTC-ACCGTATGGCTTGGATATGATGAATGCTGCTCGCTCAACACG 567
340 GCTGCATCACTGGACATACAGTATGGATAGGCTATGATGATGAATGCTGCGGACCGGT 281
568 GCCCTGGTGTATCGCAGCTGCTCGGATCCCAAGCTGCTGCGACATGCTGGCGGGG 627
280 TCCATGATCTGGCGTACGCAATGCGGTTCTTGAAGTCTCATAGATATCATAAAGCTGG 221
628 GCCCATTTGGGAGTCTCGCGGCTCTCGCTACTATTTCCATGTTGGGAACTGGGCTTAAG 687
220 GCACACTGCGGCGTCATGTTTCGGCTCAGCTTACTTCAATGACGAGCGTTGGCCAAA 161
688 GT 689
160 GT 159

RESULT 3
BI879124/c
LOCUS BI879124 502 bp mRNA linear EST 13-FEB-2002
DEFINITION fm04e08.v1 Zebrafish adult retina cDNA Danio rerio cDNA clone
IMAGE:4145367 5' similar to TR:Q9PMW4 Q9PMW4 RHODOPSIN. ; mRNA
sequence.
ACCESSION BI879124
VERSION BI879124.1 GI:16086395
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 502)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,F., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@wustl.edu
Library constructed by: Chandra Tucker and Gregory Niemi DNA
Sequencing by: Washington University Genome Sequencing Center Clome
distribution: RessourcenZentrumPrimarDatenbank, Berlin, Germany
(web address: www.rzpd.de)
Trace considered overall poor quality
Seq primer: T3 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..502
/organism="Danio rerio"
/mol_type="mRNA"
/strain="wild-type"
/db_xref="taxon:7955"
/clone="IMAGE:4145367"

```



```

/sex="mixed"
/dev_stage="1-2 years"
/lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF)'"
/clone_lib="Zebrafish adult retina cDNA"
/note="Vector: Lambda ZAP II (pBluescript SK-); Site_1:
EcoRI; Site_2: SalI; This Zebrafish library was
constructed by Dr. Susan B. Brockerhoff (email:
sbrock@u.washington.edu) RZPD library number: 760"
BASE COUNT      98 a 163 c 125 g 116 t
ORIGIN

Query Match      5.8%; Score 41.6; DB 12; Length 502;
Best Local Similarity 51.0%; Pred. No. 4.4;
Matches 98; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 300 GGACATGATCAGCACACCCCGGGTGGCGTCCCTCGTTCGGGAGAACAACTCTTCCCG 359
      |||||
Db 484 GGCCATGAGATGTGAAGCCACCATGATGGATGTTTCGCCAGCAGGAAGTTGTTG 425

QY 360 CTGCTGGGTAGGCTCACCCACGCTCGCAGTAGAGAGCCAGCGTCCCAACAGGAC 419
      |||||
Db 424 CCAGTTTGAGTCCACCATCCAGCGCTCAATGGCATGTACTCGAGCGACCATAGCCCAT 365

QY 420 AATACAGCCACGCTGATTCCTCAGCTGTTCCACATCTGCCCTCGCGCATGAGCGGT 479
      |||||
Db 364 CTCACCGCAGGTGGCTAAGTAGCCTTACAGTTGCAGCTCAGTCGGCAGAACAGGAA 305

QY 480 GCAGGACTGCAA 491
      |||||
Db 304 GTAGCGGTGCAA 293

RESULT 4
AV835132
LOCUS
DEFINITION      275 bp mRNA linear EST 22-JUN-2001
AV835132 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
spontaneum top three leaves adult, heading stage Hordeum vulgare
subsp. spontaneum cDNA clone bah24018, mRNA sequence.
ACCESSION      AV835132
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. spontaneum
Hordeum vulgare subsp. spontaneum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 275)
REFERENCE
Sato,K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished
Contact: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kazsato@rib.okayama-u.ac.jp,
URL:http://www.rib.okayama-u.ac.jp/barley/
Sato,K., Saisho,D., Takeda,K., Shini,T. and Kohara,Y. Direct
submission;
database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES
source
Location/Qualifiers
1..275
/organism="Hordeum vulgare subsp. spontaneum"
/mol_type="mRNA"
/cultivar="H602"
/db_xref="taxon:77009"
/clone="bah24018"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. spontaneum top three leaves adult, heading
stage"
30 a 123 c 67 g 46 t 9 others
BASE COUNT
ORIGIN

```

```

Query Match      5.6%; Score 40.6; DB 9; Length 275;
Best Local Similarity 48.2%; Pred. No. 6.6;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 306 GATCATGCACACCCCGGGTGGCGTCCCTCGTTCGGGAGAACAACTCTTCCCGTGGCTG 365
      |||||
Db 1 GGTCTCTGAGGAGNCCGCCCTCTGCTCCGTTGCTGCTCCGCTGCGGCTGCGGACGCGCG 60

QY 366 GGTAGCGCTCACCCCCACGCTCGCAGTAGGAAAGCGCAGCGTCCCAACAGCAATACG 425
      |||||
Db 61 CGGCCAGCGGCACACTCACTTCTCTGCTCGTTGCGCCCTCCCTCTCGGCCCCCTCG 120

QY 426 AGCCACAGTGTATCCAGCTGTTTACCATCTCGCCTCGCGGCGCATGACGCTGAGGA 485
      |||||
Db 121 ACGGCTGGGCTAGCTGGGTTGAGTTACATCTCTGCTCTCGGCGAGCGGACGCTGGCGGA 180

QY 486 CTGCAATTCTCAATCTATCCCGGCCACATACGGGTACCGGTATG 531
      |||||
Db 181 ACGNACTCGCTTCTCTCCGCCCGCGCTCGCACCGGAACG 226

RESULT 5
AV392783
LOCUS
DEFINITION      402 bp mRNA linear EST 23-APR-2002
AV392783 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CM09604_r 5', mRNA sequence.
ACCESSION      AV392783
VERSION
KEYWORDS
SOURCE
ORGANISM
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
1 (bases 1 to 402)
REFERENCE
Asamizu,E., Nakamura,Y., Sato,S., Fukuzawa,H. and Tabata,S.
A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags
DNA Res. 6 (6), 369-373 (1999)
JOURNAL
MEDLINE
PUBMED
10691129
COMMENT
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
source
Location/Qualifiers
1..402
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM09604_r"
/dev_stage="photoautotrophic growth"
/clone_lib="Chlamydomonas reinhardtii C9"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      66 a 126 c 146 g 64 t
ORIGIN

Query Match      5.6%; Score 40.6; DB 9; Length 402;
Best Local Similarity 45.3%; Pred. No. 7.5;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGTATCATTCGCTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGGCGCT 101
      |||||
Db 53 CGAGCTCATCTCGTTCATTTGTGCGCGCACTGCCAACATGAAGGACGTGCTGACGACCT 112

QY 102 GCGCATGGCTCCGGTTCTGGAGNCGCGCTGAACTATGCAACAGGGAATTTGCCGG 161
      |||||
Db 113 GGCCGCGCGCGCGAGTGGGAGGCGCGCTAGCGCACGAGTCCGTGAGTTGGGCGC 172

```


162 TTGCTCTTTCTATCTTCTCTTGGCTTTGCTGTCTGTCTGAACGGTTTCAGTTTCGGC 221
 Db 173 CCGCAAGGTTTTCACGAGATCAAGGAGTACGTTGTAACCTCAAGGCCACGAACCCAG 232
 Qy 222 TTATGAAGTCGCCAAAGTTCGGGATGTACCATGTACGAAGGACTGCTCCAATCTCAAG 281
 Db 233 CTTTCGGCGTCCGCTGGCTGGCCACATCGTTGGCGCGGCACCGCCGGCTGCTGTGAT 292
 Qy 282 CATTTGCTATGAGCGCAGCGACATGATCATGACACCCCGGCTGCGTCCCTCGGTTGCG 341
 Db 293 CCTGATGCACACGACGAGAGTTTGGCGCGCATCTACGGCGGCTGCCCATGCGCGG 352
 Qy 342 GGAGAACAACTCTTCCGCTGCTGGGT 368
 Db 353 CAGAAGACCAAGGCGACTACATGAT 379

RESULT 6
 AV392165 551 bp mRNA linear EST 23-APR-2002
 LOCUS AV392165 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
 DEFINITION cDNA clone CM083e05_r 5', mRNA sequence.
 ACCESSION AV392165
 VERSION AV392165.1 GI:6546381
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii
 ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadaceae; Chlamydomonas.

REFERENCE
 AUTHORS 1 (bases 1 to 551)
 TITLE Asamizu,E., Nakamura,Y., Sato,S., Fukuzawa,H. and Tabata,S.
 A large scale structural analysis of cDNAs in a unicellular green
 alga, Chlamydomonas reinhardtii. I. Generation of 3433
 non-redundant expressed sequence tags
 JOURNAL DNA Res. 6 (6), 369-373 (1999)
 MEDLINE 20152988
 PUBMED 10691129
 COMMENT Contact: Yasukazu Nakamura
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
 source
 1..551
 /organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="C9"
 /db_xref="taxon:3055"
 /clone="CM083e05_r"
 /dev_stage="photoautotrophic growth"
 /clone_lib="Chlamydomonas reinhardtii C9"
 /notes="Vector: pBluescriptII SK; Site_1: EcoRI; Site_2:
 XhoI"
 BASE COUNT 94 a 182 c 189 g 85 t 1 others
 ORIGIN

Query Match 5.6%; Score 40.6; DB 9; Length 551;
 Best Local Similarity 45.3%; Pred. No. 8.2;
 Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0

Qy 42 CGACCTCGTGGGTACATTCGGTCTGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCCCT 101
 Db 108 CGAGCTCATCCTGGTCATTGTGCGCGCACTGCGCAACATGAAGACGTGTGTGACGACCT 167
 Qy 102 GGCGCATGGCTCGGGTTCTGGAGGACGGGTGAACATGCAACAGGGAATTTGCCCGG 151
 Db 168 GGCGCGCGCGCGCGGATGGGAGGCGGCTACGCCACGATCGTGTAGCTTGGGCGC 227
 Qy 162 TTGCTCTTTCTATCTTCTCTTGGCTTTGCTGTCTGTCTGTGACCGTTTCAGTTTCGGC 221
 Db 228 CCGCAGAGTGTTCACGAGATCAAGGAGTACGTTGCTGAACTCAAGGCCACGAACCCAG 287
 Qy 222 TTATGAAGTCGCCAACGTGTCCGGATGTACCATGTCAACGACACTGCTCCAATCTCAAG 281

288 CTTGCCGTCGCTGCGTGGCGCCACTCGTGGCGCGGCACCGCCGGTGCCTGTCGAT 347
282 CATTGTGTATGACGAGCAGCATGATCATGCACACCCCCGGGTGGGTGCCCTTGCGTTCCG 341
348 CCGTATGACACACGACGAGAGTTTGGCGCGCATCTACGCGCGGTGCCCATGCGCGG 407
342 GGAGAACAACTCTTCCCGCTGCTGGGT 368
408 CAAGAAGAGCAAGGCGAGCTACATGAT 434

BI996341 552 bp mRNA linear EST 25-OCT-2001
1031037A07.y2 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

BI996341
BI996341.1 GI:16431115
EST.
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 552)
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
P., McDermott, J. P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Universal System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished

JOURNAL
COMMENT
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
source
1..552
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda Zap II"
/note="vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with EXase1st
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
93 a 184 c 189 g 86 t

BASE COUNT
ORIGIN

Query Match 5.6%; Score 40.6; DB 12; Length 552;
Best Local Similarity 45.3%; Pred. No. 8.2;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

Qy 42 CGACCTCGTGGGTATCATTCCTCGTCGGCGCCCCCTAGTGGCGCGCTGCACGGGCGCT 101
Db 110 CGAGCTCATCTCGGTGATGTGCGCGGCAGTCCCAACATGAAGACGTCGTGACGGACCT 169


```

/note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site 2:
ECORI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT      224 a      499 c      240 g      198 t      1 others
ORIGIN

Query Match      5.6%; Score 40.2; DB 12; Length 1162;
Best Local Similarity 54.4%; Pred. No. 13;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 313 CACACCCCCGGGTGCGTCCCTGCTGGGAGACAACTTCCCGCTGCTGGGTAGCG 372
Db 715 CCCCCCGGGTCCCTGCTGCCACACCGCGCGCCCAAAACCCCCCGGACCGCTCC 774
QY 373 CTACCCCCCAGCTCGCAGCTAGGAGCGCGAGCTCCACACGACAAATACGACCCAC 432
Db 775 CCCACTGCCACCGACACCCCCCCCCTATCGCCCCCTACCGGATCACCCTACCCAC 834
QY 433 GTCGATTCACGAGTGTTCACCATCTCGCC 461
Db 835 GCCTGATCCGGCCCTGCACACCCCGCC 863

RESULT 11
AL513886/c
LOCUS      AL513886 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA006ZG08
DEFINITION 5-PRIME. mRNA sequence.
ACCESSION  AL513886
VERSION     AL513886.2 GI:30463771
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS     Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished
COMMENT     On Feb 13, 2001 this sequence version replaced gi.12777380.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 4924.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CL0BA006ZG08P1&cluster=4924.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CL0BA006ZG08P1.

FEATURES             source
1..1201
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BA006ZG08"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched, into
double-strand cDNA was digested with Not I and cloned, into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
201 a      311 c      349 g      146 t      194 others
BASE COUNT
ORIGIN

34 GCGTTCCGCCACTCGTGGGTACATCCGCTGCTGCGCGCCCGCTAGGGGCGCTGCC 93
Db 618 GNTNTSSSSSTNNNNSSSSNNNTNTTBTBTSSSTSSSBTBTSSST 677
QY 94 AGGGCCCTGGCATGGCTCGCGTCTCGAGGACGCGTGAACATGACACAGGAAT 153
Db 678 SSSSSBBTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 737
QY 154 TTGCCCGTGTCTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 213
Db 738 TTKSBSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 797
QY 214 GCTTCGCTTAAAGTGGCAACGCTCGCGGATCACCATGTCAGCAACACTGTCTCC 273
Db 798 TBSMTSSSBTSSSSSSSBTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 854
QY 274 AACTCAAGCATCTGTATAGGCGAGCATGATCATGACACACCCCGGTCGCGCC 333
Db 855 TSSTTTNTSSYSBSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 914
QY 334 TGGTTCGGGAGAACAACTTCTCCCGCTGCTGGGTAGCGCTACCCACGCTCCGAGCT 393
Db 915 SSSSTTBSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 974
QY 394 AGNACCCAGCGTCCCGACGACAAACGACGACGCTGCGATTCACGAGCTGTACCC 453
Db 975 STTBSSTTTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 1034
QY 454 ATCTCGCTCGCGCATGAGAGCGTGCAGGATGCAATGCTCAATCTATCCCGCCAC 513
Db 1035 SNNKSSSSSSSSSSSSSSSSSSSBTTTBTSSSTSSSTSSSTSSSTSSST 1094
QY 514 ATAACGGGTACCATATGCC 533
Db 1095 SBTSSSTTTTSSSATBSB 1114

RESULT 10
BM918259
LOCUS      BM918259 1162 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6611605 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485649
5', mRNA sequence.
ACCESSION  BM918259
VERSION     BM918259.1 GI:19368638
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1162)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
            cDNA Library Prepared by: The I.M.A.G.E. Consortium (ILNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/ILNL at:
            http://image.llnl.gov
            Plate: L1CM2016 row: n column: 18
            High quality sequence stop: 567.
            Location/Qualifiers
            1..1162
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5485649"
            /tissue_type="natural killer cells, cell line"
            /lab_hosts="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_106"

FEATURES             source
1..1162
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5485649"
/tissue_type="natural killer cells, cell line"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"

```



```
Query Match      5.5%; Score 40; DB 9; Length 1201;
Best Local Similarity 26.5%; Pred. No. 15;
Matches 103; Conservative 104; Mismatches 179; Indels 3; Gaps 1;

QY 32 GCGGCTTCGCCACCTCGTGGGTACATTCCTCGCTCGCGCCCTAGGGGGGCTG 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1129 GSGCSGTGCGSBSNTYKKGKBSSSCCSVSSGSGSCSSSCSCCGGSGGGGCGGG 1070
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92 CAGGCGCCCTGGCGCATGCGTCCGGTCTCGGAGACGCGGTGAACATATGCAACAGGA 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1069 CCSCGKGGGGKSSSGSGGCCCGGG--GGSSSGSGGGGSCCWAASAVYKKGK 1013
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 152 ATTGCGCGGTGCTCTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1012 GKKTCTTTTMMAAATHTTTTWTTTTCTTAAGGGTAKVAKCCCMCCCGCCAMGCTS 953
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 CAGCTTCGCTTATGAAGTCGCGCACTGTCGCGATGTACCATGTGCAGCAAGACTGCT 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 952 GACSCSCCGCAADAACVAGGMDGAMKGTGVGSCCTTSRCKWGGGTSGMVGCATTYM 893
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 272 CCAACTCAAGCATTTGTATAGGACGAGCATGATCATGACACCCCGGGTGGCTGC 331
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 AYBSYTGTRRTWTGTSBTGTCTYASGSSGMYSSKKBGKCCMAVACSCGAGASCST 833
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 CTGGCTTCGGGAGAACACTCTTCCGCTGCTGGGTAGGCTCACCCCGCCGCTCGCAG 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 SCGSKGKTITKTTGCTGCTGAAGSABGRWTGAGGGGGGGCCCYCCMCCCCCYB 773
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 CTAGGAACGCCAGCGTCCCGCCACACACCA 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 BBCCMCWCHTKCKSKWCCRGACTYCCCCA 744
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
BJ252669/c
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 3', mRNA sequence.
ACCESSION
BJ252669
VERSION
BJ252669.1 GI:20061830
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 359)
Ogiwara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadaasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6856
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"
BASE COUNT 70 a 115 c 107 g
ORIGIN
source

Query Match      5.4%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 18;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTCGCCGACCTCGTGGGTACATTCCTCGCTCGCGCCCTAGG 83
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 CTTCAAGTCAACAGCGCGCTCTGGAAGCGCTCAGGCGGTGCAGCGCGCTCGCGTGG 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 84 GGGCGCTGCGCAGGCGCTCGCGCATGCGCTCGGGTTCTGGAGGACGCGGTGAACATG 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 GGACGCGCGCAGCGCTCGCGCATGCGCTCGCGGTGCTGCGGTGCACGTGCGCAAGG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
CC335916
LOCUS
DEFINITION
OGUJA60TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0393124,
genomic survey sequence.
ACCESSION
CC335916
VERSION
CC335916.1 GI:30805329
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 359)
Ogiwara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadaasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6856
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"
BASE COUNT 70 a 115 c 107 g
ORIGIN
source
```


Mon Dec 22 13:28:39 2003

REFERENCE 1 (bases 1 to 840)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished
COMMENT Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0393124"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

BASE COUNT 103 a 298 c 291 g 148 t
ORIGIN
Query Match 5.4%; Score 39; DB 29; Length 840;
Best Local Similarity 53.6%; Pred. No. 24;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 39 CGCGACCTCGTGGGTACATTCCGCTCGTGGCGCCCTAGGGGGCGCTGCCAGGGC 98
Db 292 CCCCGCGGTCCAGCGCCCACTCCGCTTCACTGCGCGCTCGCGGGCTCGCGGAACGC 351
Qy 99 CCTGGCGCATGGCGTCCGGGTTCTGGAGAGCGCGTGAACATATGCAACAGGGGAATTGGC 158
Db 352 TGTGGCGTCTCAAGCTGCTCTCTGGGTGCGCGCGGTAACCTCCCGACTCCGC 411
Qy 159 CGGTGGCTCTTCTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 189
Db 412 CGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 442

RESULT 15
CD446071 873 bp mRNA linear EST 03-JUN-2003
LOCUS ELO1T0207B11.b Endosperm_4 Zea mays cDNA, mRNA sequence.
DEFINITION CD446071
ACCESSION CD446071
VERSION CD446071.1 GI:31361714
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 873)
AUTHORS Lai,J., Dev,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
Messing,J.
TITLE Sequencing of the maize endosperm ESTs
JOURNAL Unpublished
COMMENT Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
Location/Qualifiers
1..873
/organism="Zea mays"
/mol_type="mRNA"

FEATURES
source
1..873
/organism="Zea mays"
/mol_type="mRNA"

/cultivar="W22"
/db_xref="taxon:4577"
/tissue_type="Endosperm of 7-23DAP"
/clone_lib="Endosperm_4"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 109 a 300 c 294 g 170 t
ORIGIN
Query Match 5.4%; Score 39; DB 14; Length 873;
Best Local Similarity 53.6%; Pred. No. 24;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 39 CGCGACCTCGTGGGTACATTCCGCTCGTGGCGCCCTAGGGGGCGCTGCCAGGGC 98
Db 583 CCCCGCGTCCAGCGCCCACTCCGCTTCACTGCGCGCTCGCGGGCTCGCGGAACGC 642
Qy 99 CCTGGCGCATGGCGTCCGGGTTCTGGAGAGCGCGTGAACATATGCAACAGGGGAATTGGC 158
Db 643 TGTGGCGTCTCAAGCTGCTCTCTGGGTGCGCGCGGTAACCTCCCGACTCCGC 702
Qy 159 CGGTGGCTCTTCTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 189
Db 703 CGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 733

Search completed: December 20, 2003, 06:54:48
Job time : 1862.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 52.1425 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-21
Perfect score: 723
Sequence: 1 ATGTTGGTAGGTATCGA.....TACTCTTTGCTCCCTAATAG 723

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	723	100.0	723	3	US-08-612-973-21
2	723	100.0	723	3	US-08-927-597-21
3	641	88.7	795	3	US-08-612-973-5
4	641	88.7	795	3	US-08-927-597-5
5	624.8	86.4	2082	3	US-08-612-973-47
6	624.8	86.4	2082	3	US-08-927-597-47
7	624.8	86.4	2433	3	US-08-612-973-49
8	624.8	86.4	2433	3	US-08-927-597-49
9	598.2	82.7	606	3	US-08-612-973-25
10	598.2	82.7	606	3	US-08-927-597-25
11	597	82.6	636	3	US-08-612-973-27
12	597	82.6	636	3	US-08-927-597-27
13	556	76.9	561	3	US-08-612-973-23
14	556	76.9	561	3	US-08-927-597-23
15	549.6	76.0	1539	2	US-08-470-426B-17
16	549.6	76.0	1863	2	US-08-470-426B-14
17	549.6	76.0	2116	3	US-08-191-160-21
18	548	75.8	932	1	US-08-081-072-15
19	548	75.8	932	1	US-08-449-093A-15
20	548	75.8	9395	3	US-09-014-416-4
21	548	74.7	9599	3	US-09-014-416-6
22	540	74.7	9472	4	US-08-150-204E-96
23	534.6	73.9	1167	2	US-08-324-977-9
24	534.6	73.9	1167	2	US-08-384-616-9
25	534.6	73.9	1167	2	US-08-904-686A-9
26	534.6	73.9	1167	3	US-09-315-850-9
27	534.6	73.9	1499	1	US-08-324-977-3

28	534.6	73.9	1499	2	US-08-384-616-3	Sequence 3, Appl
29	534.6	73.9	1499	2	US-08-904-686A-3	Sequence 3, Appl
30	534.6	73.9	1499	3	US-09-315-850-3	Sequence 3, Appl
31	534.6	73.9	6039	1	US-08-324-977-11	Sequence 11, Appl
32	534.6	73.9	6039	2	US-08-384-616-11	Sequence 11, Appl
33	534.6	73.9	6039	3	US-08-904-686A-11	Sequence 11, Appl
34	534.6	73.9	6039	3	US-09-315-850-11	Sequence 11, Appl
35	534.6	73.9	9030	1	US-08-324-977-13	Sequence 13, Appl
36	534.6	73.9	9030	2	US-08-384-616-13	Sequence 13, Appl
37	534.6	73.9	9030	2	US-08-904-686A-13	Sequence 13, Appl
38	534.6	73.9	9030	3	US-09-315-850-13	Sequence 13, Appl
39	534.6	73.9	9416	1	US-08-324-977-1	Sequence 1, Appl
40	534.6	73.9	9416	2	US-08-384-616-1	Sequence 1, Appl
41	534.6	73.9	9416	2	US-08-904-686A-1	Sequence 1, Appl
42	534.6	73.9	9416	3	US-09-315-850-1	Sequence 1, Appl
43	534.6	73.9	9416	4	US-08-823-895A-27	Sequence 27, Appl
44	493.2	68.2	742	1	US-08-081-072-18	Sequence 18, Appl
45	493.2	68.2	742	1	US-08-449-093A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-21
; Sequence 21, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..720
; FEATURE:
; NAME/KEY: mat.peptide
; LOCATION: 1..717

Mon Dec 22 13:28:37 2003

STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: US 08/612,973
APPLICATION NUMBER:
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..720
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..717
US-08-927-597-21

Query Match 100.0%; Score 723; DB 3; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.5e-183;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTCAATGATACCTTACATCGGCTTCGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAGTCAATGATACCTTACATCGGCTTCGCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCACAGGCCCTGGCGCATGGCGTCCGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCACAGGCCCTGGCGCATGGCGTCCGGTT 120
QY 121 CTGGAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTCTATCTTC 180
DB 121 CTGGAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTGTCACGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCAGCG 300
DB 241 TCCGGGATGTACCATGTGTCACGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCAGCG 300
QY 301 GACATGATATGACACACCCCGGGTGGTGGCTGGTTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATATGACACACCCCGGGTGGTGGCTGGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGGCTCGCAGCTCGCAGCTTCCGCTTATGAAGTGGCAACGTG 420
DB 361 TGCTGGGTAGCGCTCACCCCGGCTCGCAGCTCGCAGCTTCCGCTTATGAAGTGGCAACGTG 420

US-08-927-597-21
; Sequence 21, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF INVENTION: 111
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.

QY 421 ATACGACGCGCAGCTGATCCAGCTGTTCAACCATCTCGCTCGCGGATGAGACGGTG 480
DB 421 ATACGACGCGCAGCTGATCCAGCTGTTCAACCATCTCGCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATGCTCAATCTATCCCGGCCACATAACCGGTCACCGTATGGCTTGGAT 540
DB 481 CAGGACTGCAATGCTCAATCTATCCCGGCCACATAACCGGTCACCGTATGGCTTGGAT 540
QY 541 ATGATGATGAAGTGGTGGCTCAACAGCGCCCTGGTATCGAGCTGCTCCGGATCCCA 600
DB 541 ATGATGATGAAGTGGTGGCTCAACAGCGCCCTGGTATCGAGCTGCTCCGGATCCCA 600
QY 601 CAACTGCTGCTGACATGTTGGCGGGGCCCAATTTGGGAGTCTCGCGGCTCGCTAC 660
DB 601 CAACTGCTGCTGACATGTTGGCGGGGCCCAATTTGGGAGTCTCGCGGCTCGCTAC 660
QY 661 TATTCATGTTGGGGAAGTGGGCTAAGGTTTGAATGATGATGCTACTCTTTGCTCCCTAA 720
DB 661 TATTCATGTTGGGGAAGTGGGCTAAGGTTTGAATGATGATGCTACTCTTTGCTCCCTAA 720
QY 721 TAG 723
DB 721 TAG 723

RESULT 3

US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:

; NAME/KEY: mat_peptide
; LOCATION: 1..789
; US-08-612-973-5
Query Match 88.7%; Score 641; DB 3; Length 795;
Best Local Similarity 90.9%; Pred. No. 1.7e-161;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAAGGTCAATGATACCCCTTATCATGCGGCTTCCGCCACCTCGTGGGATCAATT 60
DB 1 ATGTTGGGTAAGGTCAATGATACCCCTTATCATGCGGCTTCCGCCACCTCGTGGGATCAATT 60
QY 61 CGCTCGTGGGCGCCCGCTAGGGGGCGCTCCAGGGCCCTGCGCATGCGGCTCCGGGTT 120
DB 61 CGCTCGTGGGCGCCCGCTAGGGGGCGCTCCAGGGCCCTGCGCATGCGGCTCCGGGTT 120
QY 121 CTGGAGAGCGGCGTGAACTATGCAACAGGGAAATTTGCCCGGTTGCTCTTTCTATCTTC 180
DB 121 CTGGAGAGCGGCGTGAACTATGCAACAGGGAAATTTGCCCGGTTGCTCTTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CTCTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TCCGGGATGTACCATGTGTCAGAACGAGTCTGCTCAACTCAAGCAATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACCATGTGTCAGAACGAGTCTGCTCAACTCAAGCAATTTGTATGAGGACGG 300
QY 301 GACATGATCATGACACACCCCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 GACATGATCATGACACACCCCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TGCTGGGTAGCGCTCAACCCCGCAGCTCGCAGCTAGGAAAGCCAGCGCTCCCGACAGACA 420
DB 361 TGCTGGGTAGCGCTCAACCCCGCAGCTCGCAGCTAGGAAAGCCAGCGCTCCCGACAGACA 420
QY 421 ATAGAGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438
DB 421 ATAGAGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 439 -----TCCCGAGCTGTTACCATCTCGCCCTCGCCGG 468
DB 481 GGGGACCTCTGGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 469 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 528
DB 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 600
QY 529 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 588
DB 601 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 589 CTCGGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
DB 661 CTCGGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 649 GGTCTCGCTACTATTTCCATGTTGGGAACTGCGGCTAAGGTTTGTATGATGCTACTC 708
DB 721 GGTCTCGCTACTATTTCCATGTTGGGAACTGCGGCTAAGGTTTGTATGATGCTACTC 780
QY 709 TTTGCTCCCTAATAG 723
DB 781 TTTGCTCCCTAATAG 795

RESULT 4

US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY

Mon Dec 22 13:28:37 2003

APPLICANT: BUYS, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF INVENTIONS: 111
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789

US-08-927-597-5
Query Match 88.7%; Score 641; DB 3; Length 795;
Best Local Similarity 90.9%; Pred. No. 1.7e-161;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 1 ATGTTGGTAAAGTCATCGATACCTTACATCGGCTTCGCCGACCTGCGGTACATT 60
DB 1 ATGTTGGTAAAGTCATCGATACCTTACATCGGCTTCGCCGACCTGCGGTACATT 60
QY 61 CCGCTCGTGGGCCCTTACGAGGCGCTGCGGCGCTTCGCCGACCTGCGGT 120
DB 61 CCGCTCGTGGGCCCTTACGAGGCGCTGCGGCGCTTCGCCGACCTGCGGT 120
QY 121 CTGGAGGCGGCTGAATGATGCAACAGGGAATTTGCCCGGTGCTTTCTATCTTC 180
DB 121 CTGGAGGCGGCTGAATGATGCAACAGGGAATTTGCCCGGTGCTTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CTCTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TCCGGGATGACCATGTCAGCAACGACTGCTCAAGCTCAAGCTGATGATGAGGAGCG 300
DB 241 TCCGGGATGACCATGTCAGCAACGACTGCTCAAGCTCAAGCTGATGATGAGGAGCG 300
QY 301 GACATGATCATGCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

Db 301 GACATGATCATGCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TCGTGGGTAGCGCTCACCCCAACCGCTGCGAGTACGAGCGGCGGCGGCGGCGGCGG 420
Db 361 TCGTGGGTAGCGCTCACCCCAACCGCTGCGAGTACGAGCGGCGGCGGCGGCGGCGG 420
QY 421 ATACGACGCCAGCTGAT-----TCCGAGCTGTTCCACCATCTCGCTCGCGG 438
Db 421 ATACGACGCCAGCTGAT-----TCCGAGCTGTTCCACCATCTCGCTCGCGG 480
QY 439 -----TCCGAGCTGTTCCACCATCTCGCTCGCGG 468
Db 481 GGGGACCTCTCGGATCTGCTCTCCGCTGCTCCAGCTGTTCCACCATCTCGCTCGCGG 540
QY 469 CATGAGACCGTGCAGGACTGCAATGCTCAATCTATCCGGCCACATAACGGGTACCGT 528
Db 541 CATGAGACCGTGCAGGACTGCAATGCTCAATCTATCCGGCCACATAACGGGTACCGT 600
QY 529 ATGCTTTGGGATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 588
Db 601 ATGCTTTGGGATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 589 CTCGGATCCCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
Db 661 CTCGGATCCCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 649 GGTCTGCGCTTACTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
Db 721 GGTCTGCGCTTACTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 709 TTTGCTCCCTAATAG 723
Db 781 TTTGCTCCCTAATAG 795

RESULT 5
US-08-612-973-47
Sequence 47, Application US/08612973
Patent No. 6150134
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTINOP, GUY
APPLICANT: BUYS, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs


```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-612-973-47

Query Match      86.4%; Score 624.8; DB 3; Length 2082;
Best Local Similarity 90.5%; Pred. No. 4.7e-157;
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;

QY 4 TTGGGTAAGGTATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATTCCG 63
Db 4 TTGGGTAAGGTATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATTCCG 63
QY 64 CTCGTCCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123
Db 64 CTCGTCCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123
QY 124 GAGGACGGCGTGAATATCAACAGAGGAATTTGCCCGGTTGCTCTTCTATCTTCCTC 183
Db 124 GAGGACGGCGTGAATATCAACAGAGGAATTTGCCCGGTTGCTCTTCTATCTTCCTC 183
QY 184 TTGGCTTTGCTGCTGCTGTGACCGTTCCAGCTTCGCTTATGAAGTGGCGACGTGTC 243
Db 184 TTGGCTTTGCTGCTGCTGTGACCGTTCCAGCTTCGCTTATGAAGTGGCGACGTGTC 243
QY 244 GGGATGTACATGTCTACGAAGACTCTCAACTCAAGCATTTGTATGAGGAGCGGAC 303
Db 244 GGGATGTACATGTCTACGAAGACTCTCAACTCAAGCATTTGTATGAGGAGCGGAC 303
QY 304 ATGATCATGCACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCGCTGC 363
Db 304 ATGATCATGCACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCGCTGC 363
QY 364 TGGGTAGCGCTACCCCGAGCTCGAGTAGGAAGCGCAGCTGCCACAGCAATA 423
Db 364 TGGGTAGCGCTACCCCGAGCTCGAGTAGGAAGCGCAGCTGCCACAGCAATA 423
QY 424 CGACGCCAGCTCGAT----- 438
Db 424 CGACGCCAGCTCGATTTGCTGTTGGGGGGGGTGTCTTCTGTTCCGCTATGATAGTGGGG 483
QY 439 -----TCCAGCTGTTCAACATCTCGCTCGCCGGCAT 471
Db 484 GACCTCTGGGATCTGCTTCTCTGCTCTCCAGCTGTTCAACATCTCGCTCGCCGGCAT 543
QY 472 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTACCGTATG 531
Db 544 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTACCGTATG 603
QY 532 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
Db 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
QY 592 CGGATCCCAAGCTGCTGTTGACATGTTGGCGGGGGCCCATTTGGGAGTCTCTGGCGGGT 651
Db 664 CGGATCCCAAGCTGCTGTTGACATGTTGGCGGGGGCCCATTTGGGAGTCTCTGGCGGG 723
QY 652 CTCGCTACTATTCCATGTTGGGGAACCTGGGCTTAAGGTTTGTGATGCTACTCTTT 711
Db 724 CTCGCTACTATTCCATGTTGGGGAACCTGGGCTTAAGGTTTGTGATGCTACTCTTT 783
QY 712 GC 713
Db 784 GC 785
```

RESULT 6

```
US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-927-597-47
```

```
Query Match      86.4%; Score 624.8; DB 3; Length 2082;
Best Local Similarity 90.5%; Pred. No. 4.7e-157;
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;

QY 4 TTGGGTAAGGTATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATTCCG 63
Db 4 TTGGGTAAGGTATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATTCCG 63
QY 64 CTCGTCCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123
Db 64 CTCGTCCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123
QY 124 GAGGACGGCGTGAATATCAACAGAGGAATTTGCCCGGTTGCTCTTCTATCTTCCTC 183
Db 124 GAGGACGGCGTGAATATCAACAGAGGAATTTGCCCGGTTGCTCTTCTATCTTCCTC 183
```


Db 1015 CGGATCCCAAGCTGTCTGGACATGGTGGCGGGGCCCAATTGGGGAGTCTCTGGCGGC 1074
QY 652 CTCGCTACTATTCCATGGTGGGGAACCTGGGCTTAAGGTTTTCATTTGATGCTACTCTTT 711
Db 1075 CTCGCTACTATTCCATGGTGGGGAACCTGGGCTTAAGGTTTTCATTTGATGCTACTCTTT 1134
QY 712 GC 713
Db 1135 GC 1136

RESULT 8

US-08-927-597-49
; Sequence 49, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2430
; FEATURE:
; NAME/KEY: mat.peptide
; LOCATION: 1..2427
; US-08-927-597-49

Query Match 86.4%; Score 624.8; DB 3; Length 2433;
Best Local Similarity 90.5%; Pred. No. 4.9e-157;
Matches 708; Conservative 2; Indels 72; Gaps 1;
QY 4 TTGGGTAAGTCAATCCCTTACATGCGGCTTCGCGGACCTCGGGGTACATCCG 63
|||||

Db 355 TTGGGTAAGTCAATCCCTTACATGCGGCTTCGCGGACCTCGTGGGGTACATCCG 414
QY 64 CTCGTGGCGGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCCGGGTTCTG 123
Db 415 CTCGTGGCGGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCCGGGTTCTG 474
QY 124 GAGGACGGGTGAACCTATGCAACAGGGAATTTGCCCGGTGCTCTTTCTCTATCTTCCTC 183
Db 475 GAGGACGGGTGAACCTATGCAACAGGGAATTTGCCCGGTGCTCTTTCTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTCTGACCGCTTCAGCGTTCCGCTTATGAAGTCGCGCAACGTTGCC 243
Db 535 TTGGCTTTGCTGCTCTGACCGCTTCAGCGTTCCGCTTATGAAGTCGCGCAACGTTGCC 594
QY 244 GGGATGACCATGTGTACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGCAGGGAC 303
Db 595 GGGATGACCATGTGTACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGCAGGGAC 654
QY 304 ATGATCATGCACACCCCGGGTGGCTGCTGGTTCGGGAGAAACAACTCTTCCCGCTGC 363
Db 655 ATGATCATGCACACCCCGGGTGGCTGCTGGTTCGGGAGAAACAACTCTTCCCGCTGC 714
QY 364 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAACGCCAGCGTCCCAACCCAGCAATA 423
Db 715 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAACGCCAGCGTCCCAACCCAGCAATA 774
QY 424 CGAGCCACGTCGAT----- 438
Db 775 CGAGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 834
QY 439 -----TCCAGCTGTTCAACATCTCGCTCGCGGCGCAT 471
Db 835 GACCTCTGGGATGCTGCTTCTCGTCTCCAGCTGTTCAACATCTCGCTCGCGGCGCAT 894
QY 472 GAGAGGTGCAGGACTGCAATTCCTCAATCTATCCGCGCACATAACGGGTACCCGTATG 531
Db 895 GAGAGGTGCAGGACTGCAATTCCTCAATCTATCCGCGCACATAACGGGTACCCGTATG 954
QY 532 GCTTGGGATATGATGATGAACCTGCTCGCTCAAAACGGCCCTGCTGATCGAGCTGCTC 591
Db 955 GCTTGGGATATGATGATGAACCTGCTCGCTCAAAACGGCCCTGCTGATCGAGCTGCTC 1014
QY 592 CGGATCCCAAGCTGCTGTGACATGTTGGCGGGGGCCCATTTGGGAGTCTCTGGCGGT 651
Db 1015 CGGATCCCAAGCTGCTGTGACATGTTGGCGGGGGCCCATTTGGGAGTCTCTGGCGGC 1074
QY 652 CTCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTTCATTTGATGCTACTCTTT 711
Db 1075 CTCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTTCATTTGATGCTACTCTTT 1134
QY 712 GC 713
Db 1135 GC 1136

RESULT 9

US-08-612-973-25
; Sequence 25, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.

us-09-899-303a-21.rn1

Mon Dec 22 13:28:37 2003

ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..603
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..600
US-08-612-973-25

Query Match 82.7%; Score 598.2; DB 3; Length 606;
Best Local Similarity 99.5%; Pred. No. 3.8e-150; Indels 0; Gaps 0;
Matches 600; Conservative 0; Mismatches 3;
QY 1 ATGTTGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTCGCGGCCCCCTTAGGGGGCGCTCCAGAGGCCCTGGCGCATGCCGTCGGGTT 120
DB 61 CCGCTCGTCGCGGCCCCCTTAGGGGGCGCTCCAGAGGCCCTGGCGCATGCCGTCGGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGTGACCGTTCCAGCTTCGGTTATGAAGTGGCAAGTG 240
DB 181 CTCTTGGCTTTGCTGCTGCTGTGACCGTTCCAGCTTCGGTTATGAAGTGGCAAGTG 240
QY 241 TCCGGATGTACCATGTACGAACACACTGCTCCAACTCAAGCATTTGTTATGAGCGAGCG 300
DB 241 TCCGGATGTACCATGTACGAACACACTGCTCCAACTCAAGCATTTGTTATGAGCGAGCG 300
QY 301 GACATGATCATGACACACCCCGGGTTCGGTTCGCTTCGGTGGGAGAAACAATCTTCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTTCGGTTCGCTTCGGTGGGAGAAACAATCTTCCGC 360
QY 361 TGTGGGTAGCGTCAACCCCGAGCTCCGAGCTAGGAAGCCGAGGTCGCCACACGACA 420
DB 361 TGTGGGTAGCGTCAACCCCGAGCTCCGAGCTAGGAAGCCGAGGTCGCCACACGACA 420
QY 421 ATACGAGCGCACGTGCATTTCCAGCTGTTTCCACATCTCGCTTCGCGGATGAGCGGTG 480
DB 421 ATACGAGCGCACGTGCATTTCCAGCTGTTTCCACATCTCGCTTCGCGGATGAGCGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCGGGCAATAACGGGTCCAGCTATGCTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCGGGCAATAACGGGTCCAGCTATGCTTGGGAT 540

QY 541 ATGATGATGAACCTGGTCGCTCAACAGGCCCTGGTGGTATCGAGCTGCTCCGGATCCCA 600
DB 541 ATGATGATGAACCTGGTCGCTCAACAGGCCCTGGTGGTATCGAGCTGCTCCGGATCCCTC 600
QY 601 CAA 603
DB 601 TAA 603

RESULT 10
US-08-927-597-25
Sequence 25, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..603
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..600
US-08-927-597-25

Query Match 82.7%; Score 598.2; DB 3; Length 606;
Best Local Similarity 99.5%; Pred. No. 3.8e-150; Indels 0; Gaps 0;
Matches 600; Conservative 0; Mismatches 3;
QY 1 ATGTTGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4000
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..633
NAME/KEY: mat_peptide
LOCATION: 1..630
US-08-927-597-27

Query Match 82.6%; Score 597; DB 3; Length 636;
Best Local Similarity 100.0%; Pred. No. 8.2e-150;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60
QY 61 CGCTCGTCCGCGCCCTAGGGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGTT 120
DB 61 CGCTCGTCCGCGCCCTAGGGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTGTCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
DB 181 CTCTTGGCTTGTCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
QY 241 TCCGGATGTACCATGTCAACAGCACTGCTCCAACTCAAGCATTTGTATGAGCAGCG 300
DB 241 TCCGGATGTACCATGTCAACAGCACTGCTCCAACTCAAGCATTTGTATGAGCAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGCTGCTCGCTTCCGGAGAACAACTCTTCCCGC 360

DB 301 GACATGATCATGCACACCCCGGGTGGCTGCTCGCTTCCGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTACCCCGACGCTCCAGCTAGGAACGCGCAGCTCCCGACCAAGACA 420
DB 361 TGCTGGGTAGCGCTACCCCGACGCTCCAGCTAGGAACGCGCAGCTCCCGACCAAGACA 420
QY 421 ATAGGACGCGCAGCTCGATTCCCGAGCTGTTTACCATCTCGCTCGCGGCATGAGACGGTG 480
DB 421 ATAGGACGCGCAGCTCGATTCCCGAGCTGTTTACCATCTCGCTCGCGGCATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGACATACCGGTCCAGTATGCTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGACATACCGGTCCAGTATGCTTGGGAT 540
QY 541 ATGATGATGAACGTGCTCGCTACCAAGCGCCCTGGTGTATCGAGCTGCTCGGATC 597
DB 541 ATGATGATGAACGTGCTCGCTACCAAGCGCCCTGGTGTATCGAGCTGCTCGGATC 597

RESULT 13
US-08-612-973-23
; Sequence 23, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; NAME/KEY: mat_peptide
; LOCATION: 1..555
US-08-612-973-23

Query Match 76.9%; Score 556; DB 3; Length 561;
Best Local Similarity 100.0%; Pred. No. 6.4e-139;

Matches	556;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	ATGTTGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCCGACCTGTCGGGTACATT	60						
Db	1	ATGTTGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCCGACCTGTCGGGTACATT	60						
Qy	61	CCGCTCGTCGGCGCCCTTCATAGGGGGCGCTGCCAAGGGCCCTGGCGCATGGCGTCCGGGT	120						
Db	61	CCGCTCGTCGGCGCCCTTCATAGGGGGCGCTGCCAAGGGCCCTGGCGCATGGCGTCCGGGT	120						
Qy	121	CTGAGGACCGCGTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC	180						
Db	121	CTGAGGACCGCGTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC	180						
Qy	181	CTCTTGCTTTGCTGTCTGCTGTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTG	240						
Db	181	CTCTTGCTTTGCTGTCTGCTGTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTG	240						
Qy	241	TCGGGATGTACCATGTTCAGAAAGACTGTCCTCAACTCAAGCATGTGTATGAGGCACGG	300						
Db	241	TCGGGATGTACCATGTTCAGAAAGACTGTCCTCAACTCAAGCATGTGTATGAGGCACGG	300						
Qy	301	GACATGATCATGCACACCCCGGCTGGCTGCGTTCCGGAGAACAACTCTTCCCGC	360						
Db	301	GACATGATCATGCACACCCCGGCTGGCTGCGTTCCGGAGAACAACTCTTCCCGC	360						
Qy	361	TGCTGGGTAGCGTCAACCCCAACGCTGCAAGTCTAGGAAACGCGGTCCCAACACGACA	420						
Db	361	TGCTGGGTAGCGTCAACCCCAACGCTGCAAGTCTAGGAAACGCGGTCCCAACACGACA	420						
Qy	421	ATACGACGCCACGTCGATTCACCATCTCCAGCTGTTACCATCTCGCCTCGCCGCGATGAGACGGTG	480						
Db	421	ATACGACGCCACGTCGATTCACCATCTCCAGCTGTTACCATCTCGCCTCGCCGCGATGAGACGGTG	480						
Qy	481	CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGGAT	540						
Db	481	CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGGAT	540						
Qy	541	ATGATGATGAACCTGGT	556						
Db	541	ATGATGATGAACCTGGT	556						

RESULT 14
US-08-927-597-23
; Sequence 23, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P. C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973

RESULT 15
US-08-470-426B-17
; Sequence 17, Application US/08470426B

Patent No. 5856458
GENERAL INFORMATION:
APPLICANT: Okamoto, Hiroaki
APPLICANT: Nakamura, Tetsuo
TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B-
TITLE OF INVENTION: HEPATITIS VIRUS
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beverage, DeGrandi, Wellacher & Young.
ADDRESSEE: L.L.P.
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,426B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-153402
FILING DATE: 12-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Wellacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/59-47083.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1539 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HIS-08-470-426B-17

Query Match	76.0%;	Score 549.6;	DB 2;	Length 1539;
Best Local Similarity	84.5%;	Pred. No. 4.4e-137;		
Matches 661;	Conservative 0;	Mismatches 49;	Indels 72;	Gaps 1;
QY	4	TTGGGTAAAGGTCATCGATACCTTACATGCGGCTTGCCGACCTCGTGGGGTACATTCGG	63	
DB	355	TTGGGTAAAGGTCATCGATACCTTACATGCGCTTGCCGATCTCATGGGTATATTCCTC	414	
QY	64	CTCGTCGGCGCCCCCTTAGGGGGGCGCTGCCAGGSCCTCGCGCATGGCGTCCGGGTTCTG	123	
DB	415	CTCGTCGGCGCCCCCTTAGGGGGGCGCTGCCAGGSCCTTGGCAACAGGTGTCGGGGTCTG	474	
QY	124	GAGGACGGCGTGAATATGTGCAACAGGNAATTTGCCCGGTTGCTTTCTATCTTCCTC	183	
DB	475	GAGGACGGCGTGAATATGTGCAACAGGGAACATTGCCCGGTTGCTTTCTATCTTCCTC	534	
QY	184	TTGGCTTTGCTGTCCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGGCGCAACGFTGCC	243	
DB	535	TTGGCTTTGCTGTCCTGTTTGACCATCCAGCTTCCGCTTATGAAGTGGCGCAACGFTGCC	594	
QY	244	GGGATGTACCATGTGTCAGAAAGCATGTGCTCCAACTCAAGCATTTGTATGAGGCAGCGAAC	303	
DB	595	GGGATATACCATGTGTCAGAAAGCATGTGCTCCAACTCAAGCATTTGTATGAGGCAGCGAAC	654	
QY	304	ATGATCATGCACACCCCGGGTGGCTGCCCTGCGTTGGGAGAACAACTCTTCCCGCTCG	363	
DB	655	ATGATCATGCATCTACTCCCGGGTGGCTGCCCTGCGTTGGGAGGAACAACAGCTCCCGCTTG	714	
QY	364	TGGGTAGCGCTCACCCCCACGCTCGGAGTAGGAAGCCAGCGTCCCGACACGACATA	423	

D _b	715	TGGGTAGCGCTCACTCCACGCTCGCGGCAGGAATGCCAGGGTCCCCACTACGAAATA	774
Q _y	424	CGACGCCACGTCCA-----	437
D _b	775	CGACGCCACGTCACTTGCTGTGGGGGGCGGTGCTTTCTGTCGCGTATGTACGTGGGG	834
Q _y	438	-----TTCCAGCTGTTCCACCATCTGCGCTCGCGGCAT	471
D _b	835	GATCTCTGGCGGATCTGTTTTCTCGTCTCCAGCTGTTCACCTTCTCGCGCTCGCGGCAT	894
Q _y	472	GAGACGGTGAGGATGCAATTGCTCAATTATCCCGGCCACATAACGGGTCAACCGTATG	531
D _b	895	GAGACATGCGAGACTGCAACTGCTCAATCTATCCCGGCCAATTATCAGGTCAACCGCATG	954
Q _y	532	GCTTGGGATATGATGAATGTCGTCCTACAAACGGCCCTGCTGCTATCGCAGCTGCCTC	591
D _b	955	GCTTGGGATATGATGATGAATGTCACCTACCAACAGCCCTAGTGTGTGTCGCAAGTTGCTC	1014
Q _y	592	CGGATCCCACAAGCTGTCGTGGACATGGTGGCGGGGGCCCAATTGGGAGTCTCTGGCGGGT	651
D _b	1015	CGGATCCCACAAGCTGTCGTGGACATGGTGGCGGGGGGCCACTGGGAGTCTCTGGCGGGC	1074
Q _y	652	CTCGCTACTATTCCATGCTGGGAACTGGGCTTAAGGTTTTGATTGTGATGCTACTCTTT	711
D _b	1075	CTTGCTACTATTCCATGGTAGGAACTGGGCTAAGGTCCTGATTGTGGCGCTACTCTTC	1134
Q _y	712	GC	713
D _b	1135	GC	1136

Search completed: December 20, 2003, 07:03:07
Job time : 56.1425 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2124.1 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-23

Perfect score: 561

Sequence: 1 ATGTTGGGTAAAGTCATCGA.....TCGATGTAAGTCTGTAATAG 561

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	561	100.0	561	6	A48685 Sequence 23
2	561	100.0	561	6	AR157337 Sequence
3	561	100.0	561	6	AX452772 Sequence
4	561	100.0	561	6	AX685024 Sequence
5	556	99.1	606	6	A48687 Sequence 25
6	556	99.1	606	6	AR157338 Sequence
7	556	99.1	606	6	AX452774 Sequence
8	556	99.1	606	6	AX685026 Sequence
9	556	99.1	636	6	A48689 Sequence 27
10	556	99.1	636	6	AR157339 Sequence
11	556	99.1	636	6	AX452776 Sequence
12	556	99.1	636	6	AX685028 Sequence
13	556	99.1	723	6	A48683 Sequence 21
14	556	99.1	723	6	AR157336 Sequence
15	556	99.1	723	6	AX452770 Sequence
16	556	99.1	723	6	AX685022 Sequence
17	474	84.5	795	6	A48667 Sequence 5
18	474	84.5	795	6	AR157325 Sequence
19	474	84.5	795	6	AX452754 Sequence
20	474	84.5	795	6	AX685006 Sequence
21	471	84.0	2082	6	A48709 Sequence 47
22	471	84.0	2082	6	AR157350 Sequence
23	471	84.0	2082	6	AX452796 Sequence
24	471	84.0	2082	6	AX685048 Sequence
25	471	84.0	2433	6	A48711 Sequence 49
26	471	84.0	2433	6	AR157351 Sequence
27	471	84.0	2433	6	AX452798 Sequence
28	471	84.0	2433	6	AX685050 Sequence
29	461.4	82.2	633	6	A48669 Sequence 7
30	461.4	82.2	633	6	AR157326 Sequence
31	461.4	82.2	633	6	AX452756 Sequence
32	461.4	82.2	633	6	AX685008 Sequence
33	454	80.9	636	6	A48675 Sequence 13
34	454	80.9	636	6	AR157329 Sequence
35	454	80.9	636	6	AX452762 Sequence
36	454	80.9	636	6	AX685014 Sequence
37	426.6	76.0	9379	14	AF165052 Hepatitis
38	426.6	76.0	9379	14	AF207766 Hepatitis
39	426.6	76.0	9379	14	HCVPOLYP Hepatitis
40	423.4	75.5	3296	14	AB008441 Hepatitis
41	423.4	75.5	3296	14	AB008442 Hepatitis
42	423.4	75.5	3296	14	AB008443 Hepatitis
43	423.4	75.5	9344	14	AB049096 Hepatitis
44	423.4	75.5	9379	14	AF165051 Hepatitis
45	421.8	75.2	9359	14	AF313916 Hepatitis

ALIGNMENTS

RESULT 1
A48685
LOCUS A48685 Sequence 23 from Patent WO9604385. 561 bp DNA linear PAT 07-MAR-1997
DEFINITION A48685
ACCESSION A48685
VERSION A48685.1 GI:2302398
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 561)
AUTHORS Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 23 15-FEB-1996;

INNOGENETICS NV (BE)	Other publication CA 2172273 960215
COMMENT	Other publication AU 3382495 960304.
FEATURES	Location/Qualifiers
source	1. .561
	/organism="unidentified"
	/mol_type="genomic DNA"
	/db_xref="taxon:32644"
CDS	1. .558
	/notes="unnamed protein product"
	/codon_start=1
	/protein_id="CAA03135.1"
	/db_xref="GI:2302399"
	/translation="MLGKVIDLTTCGFADLVGYIPLVGAPLGGAARALAHGVRVLBDG VNTATGNLPCGSFSIFILLALISCLTVPASAVEVRNVSGMVHVTNDCSNSIIVYEAAAD IMHTPGGCPVCVRNNSSRCWALVTLPAARNASVPTTIRRHVDSQLFTTISPRRHETV QDNCNSIYPGHITGHRMADWMNNW"
mat_peptide	1. .555
	/product="unnamed"
BASE COUNT	103 a 176 c 155 g 127 t
ORIGIN	
Query Match	100.0%; Score 561; DB 6; Length 561;
Best Local Similarity	100.0%; Pred. No. 2e-118;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGTTGGGTAGGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
Db	1 ATGTTGGGTAAAGGTATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
QY	61 CGCGTCGTCGGCGCCCGCCCTAGAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
Db	61 CGCGTCGTCGGCGCCCGCCCTAGAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
QY	121 CTGGAGGACGGGTGAATGATGCAACAGGGAAATTGCGCGGTTGCTCTTCTCTATCTTC 180
Db	121 CTGGAGGACGGCGTGAATGATGCAACAGGGAAATTGCGCGGTTGCTCTTCTCTATCTTC 180
QY	181 CTCTTGGCTTTGCTGCTCTGTCTGACCGTTCACCGTTCCAGCTTTCCGCTTTATGAA GTGCGCAACGTG 240
Db	181 CTCTTGGCTTTGCTGCTCTGTCTGACCGTTCACCGTTCCAGCTTTCCGCTTTATGAA GTGCGCAACGTG 240
QY	241 TCCGGGATGTACATGTATCGAAACGACTGTCTCAGCTCAAGCAATGTGTATGAGGCAGCG 300
Db	241 TCCGGGATGTACCATGTATCGAAACGACTGTCTCAGCTCAAGCAATGTGTATGAGGCAGCG 300
QY	301 GACATGATCATGCACACCCCGGGTGGCGCTTCGCTTCGAGAGAACAACTCTTCCCGC 360
Db	301 GACATGATCATGCACACCCCGGGTGGCGCTTCGCTTCGAGAGAACAACTCTTCCCGC 360
QY	361 TGCTGGGTAGCGTCAACCCCAAGCTGCGAGCTAGGAACGCCAGCGTCCCACCAACGACA 420
Db	361 TGCTGGGTAGCGTCAACCCCAAGCTGCGAGCTAGGAACGCCAGCGTCCCACCAACGACA 420
QY	421 ATAGCAGCCACGTCGATTCCCAGCTGTTACCACTCTCGCTCCCGCGCATGAGACGGTG 480
Db	421 ATAGCAGCCACGTCGATTCCCAGCTGTTACCACTCTCGCTCCCGCGCATGAGACGGTG 480
QY	481 CAGGACTGCAATTTGCTCAATCTATCTCCCGGCCACATAACGGGTCAACGGTTGGGTAT 540
Db	481 CAGGACTGCAATTTGCTCAATCTATCTCCCGGCCACATAACGGGTCAACGGTTGGGTAT 540
QY	541 ATGATGATGAACCTGGTAAATAG 561
Db	541 ATGATGATGAACCTGGTAAATAG 561
RESULT 2	
AR157337	AR157337
LOCUS	Sequence 23 from patent US 6245503.
DEFINITION	AR157337
ACCESSION	AR157337.1
VERSION	GI:16218270
	linear PAT 17-OCT-2000

TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 23 05-JUN-2002;
FEATURES Innogenetics N.V. (BE)
source Location/Qualifiers
1..561
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..558
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38073.1"
/db_xref="GI:21712458"
/translation="MLGKVIDLTLCGFADLVGYIPLVGAPLGGAARALAHGVRLDGG
VNYATGNLPGCSFSLIALALLSLCTVPASAYEVNRNMGVHYVNTDCNSNLSIVVEAADM
IMHTPCVPCVRENNSRCWALPTLAARNASVPTTTIRRHVDSQLFTTISPRRHETV
QDCNCSYIPGHITGHRMADMMMNW"
CDS 1..555
mat_peptide 103 a 176 c 155 g 127 t
BASE COUNT 103 a 176 c 155 g 127 t
ORIGIN
Query Match 100.0%; Score 561; DB 6; Length 561;
Best Local Similarity 100.0%; Pred. No. 2e-118;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTGGGTAAGTCAATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAGTCAATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60
Qy 61 CCGCTCGTGGGCGCCCTTAGGGGGCGTGCAGGGCCCTGGCGCATGGCGTTCGGGTT 120
Db 61 CCGCTCGTGGGCGCCCTTAGGGGGCGTGCAGGGCCCTGGCGCATGGCGTTCGGGTT 120
Qy 121 CTGGAGGACGGCTGAATATGCAACAGGGAATTTGCCGGTGTCTCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGCTGAATATGCAACAGGGAATTTGCCGGTGTCTCTTCTCTATCTTC 180
Qy 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAACGTG 240
Db 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAACGTG 240
Qy 241 TCGGGATGTACATGTACGAACGACTGCTCCAACTCAAGCATTTGTATGAGCGACGG 300
Db 241 TCGGGATGTACATGTACGAACGACTGCTCCAACTCAAGCATTTGTATGAGCGACGG 300
Qy 301 GACATCATATGCACACCCCGGGTGGCGCTTCGCTTCGCGGAGAACAACTTCCCGC 360
Db 301 GACATCATATGCACACCCCGGGTGGCGCTTCGCTTCGCGGAGAACAACTTCCCGC 360
Qy 361 TGCTGGGTAGCGTCAACCCCGCGTTCGAGTACGAGTACGAGTACGAGTACGAGTACG 420
Db 361 TGCTGGGTAGCGTCAACCCCGCGTTCGAGTACGAGTACGAGTACGAGTACGAGTACG 420
Qy 421 ATACGACGCAACGTCATCTATCCCGGCGACATACCGGTACCGGTATGGGTGGGAT 540
Db 421 ATACGACGCAACGTCATCTATCCCGGCGACATACCGGTACCGGTATGGGTGGGAT 540
Qy 541 ATGATGATGAACCTGGTAATAG 561
Db 541 ATGATGATGAACCTGGTAATAG 561

RESULT 4
AX685024 AX685024 561 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 23 from Patent WO02055548.
ACCESSION AX685024
VERSION AX685024.1 GI:29371429

RESULT	4
AX685024	
LOCUS	AX685024 .
DEFINITION	Sequence 23 from Patent WO02055548.
ACCESSION	AX685024
VERSION	AX685024.1 GI:29371429
	561 bp DNA linear PAT 29-MAR-2003

Db 421 ATACGAGCCACGTCGATTCCCGATTTCACCATCTCGCTCGCGGATGAGACGGTG 480

Qy 481 CAGGACTGCAATTCGTCATCTATCCCGCCACATACGCGTCAACGCTATGCTTGGGAT 540

Db 481 CAGGACTGCAATTCGTCATCTATCCCGCCACATACGCGTCAACGCTATGCTTGGGAT 540

Qy 541 ATGATGATGAACCTGGT 556

Db 541 ATGATGATGAACCTGGT 556

RESULT 6

LOCUS ARL157338 606 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 25 from patent US 6245503.

ACCESSION ARL157338

VERSION ARL157338.1 GI:16218271

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 606)

AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.

TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use

JOURNAL Patent: US 6245503-A 25 12-JUN-2001;

FEATURES

source Location/Qualifiers

1..606

/organism="unknown"

BASE COUNT 109 a 193 c 167 g 137 t

ORIGIN

Query Match 99.1%; Score 556; DB 6; Length 606;

Best Local Similarity 100.0%; Pred. No. 2.8e-117;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGGACCTCGTGGGTACATT 60

Qy 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTT 120

Db 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTT 120

Qy 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Qy 181 CTCTTGGCTTTGCTGTCCTGCTGACCGCTTCAGCGTTCGCTTATGAAAGTGGCAACGTT 240

Db 181 CTCTTGGCTTTGCTGTCCTGCTGACCGTTCAGCGTTCGCTTATGAAAGTGGCAACGTT 240

Qy 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCG 300

Db 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCG 300

Qy 301 GACATGATCATGACACACCCCGGTCGCTCGCTTCGCTTCGCGGAGAACTCTTCCCGC 360

Db 301 GACATGATCATGACACACCCCGGTCGCTCGCTTCGCTTCGCGGAGAACTCTTCCCGC 360

Qy 361 TGCTGGGTAGCGTCAACCCCGGTCGCTCGCTTCGCTTCGCGGAGAACTCTTCCCGC 420

Db 361 TGCTGGGTAGCGTCAACCCCGGTCGCTCGCTTCGCTTCGCGGAGAACTCTTCCCGC 420

Qy 421 ATACGAGCCACGTCGATTCCCGATTTCACCATCTCGCTCGCGGATGAGACGGTG 480

Db 421 ATACGAGCCACGTCGATTCCCGATTTCACCATCTCGCTCGCGGATGAGACGGTG 480

Qy 481 CAGGACTGCAATTCGTCATCTATCCCGCCACATACGCGTCAACGCTATGCTTGGGAT 540

Db 481 CAGGACTGCAATTCGTCATCTATCCCGCCACATACGCGTCAACGCTATGCTTGGGAT 540

Qy 541 ATGATGATGAACCTGGT 556

RESULT 5

LOCUS A48687 606 bp DNA linear PAT 07-MAR-1997

DEFINITION Sequence 25 from Patent WO9604385.

ACCESSION A48687

VERSION A48687.1 GI:2302400

KEYWORDS

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 606)

AUTHORS Maertens,G., Bosman,F., De M.G. and Buyse,M.

TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

JOURNAL Patent: WO 9604385-A 25 15-FEB-1996;

COMMENT INNOCENTICS NV (BE)

Other publication CA 2172273 960215

Other publication AU 3382495 960304.

FEATURES

source Location/Qualifiers

1..606

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

1..603

/note="unnamed protein product"

/codon_start=1

/protein_id="CAA03136.1"

/db_xref="GI:2302401"

/translation="MLGKVIDTLTCGFADLVGPIPLVGAPLGARALAHGVRLVDG VNYATGNLPQCSFISFILLALLSLCTWPAASAYEVNYSVMHYVNDSCNSIYVRAADM IMPTGCVPCVRNNSSRCWVALTPTLAARNASVPTTIRRHVDLSQFTISPRRHETV QDCNSIYPGHITGHRMWDMMWNSPTTALVVSQLLRIL"

mat_peptide 1..600

/product="unnamed"

BASE COUNT 109 a 193 c 167 g 137 t

ORIGIN

Query Match 99.1%; Score 556; DB 6; Length 606;

Best Local Similarity 100.0%; Pred. No. 2.8e-117;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCGGACCTCGTGGGTACATT 60

Qy 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTT 120

Db 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTT 120

Qy 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Qy 181 CTCTTGGCTTTGCTGTCCTGCTGACCGTTCAGCGTTCGCTTATGAAAGTGGCAACGTT 240

Db 181 CTCTTGGCTTTGCTGTCCTGCTGACCGTTCAGCGTTCGCTTATGAAAGTGGCAACGTT 240

Qy 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCG 300

Db 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCG 300

Qy 301 GACATGATCATGACACCCCGGTCGCTCGCTTCGCTTCGCGGAGAACTCTTCCCGC 360

Db 301 GACATGATCATGACACCCCGGTCGCTCGCTTCGCTTCGCGGAGAACTCTTCCCGC 360

Qy 361 TGCTGGGTAGCGTCAACCCCGGTCGCTCGCTTCGCTTCGCGGAGAACTCTTCCCGC 420

Db 361 TGCTGGGTAGCGTCAACCCCGGTCGCTCGCTTCGCTTCGCGGAGAACTCTTCCCGC 420

Qy 421 ATACGAGCCACGTCGATTCCCGATTTCACCATCTCGCTCGCGGATGAGACGGTG 480


```
Db 541 ATGATGATGAAGTGGT 556
|||||
RESULT 7
AX452774
LOCUS AX452774 606 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 25 from Patent EP1211315.
ACCESSION AX452774
VERSION AX452774.1 GI:21712459
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 25 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..606
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..603
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38074.1"
/db_xref="GI:21712460"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLDGG
VNYATGNLPGCSFSIFLLALLSCLTVPASAYEVNVMYHVTNDCSNSIVVEAADM
INHTGCVPCVRENSSRCWVALPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIVPGHITGHRMAMWMMNWSPTTALVVSQLLRL"
mat_peptide 1..600
BASE COUNT 109 a 193 c 167 g 137 t
ORIGIN
Query Match 99.1%; Score 556; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTGCATCATACCCCTTACATCGGCTTCGCCACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAGTGCATCATACCCCTTACATCGGCTTCGCCACCTCGTGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTTGGCGCATGGCTCCGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTTGGCGCATGGCTCCGGTT 120
QY 121 CTGGAGGAGCGGCTGAACCTATGCAACAGGGAATTTGCCGGTGTCTTTCTATCTTC 180
Db 121 CTGGAGGAGCGGCTGAACCTATGCAACAGGGAATTTGCCGGTGTCTTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
Db 181 CTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
QY 241 TCGGGATGATACATGTCAGAACGACTGCTCAACTCAAGCATTTGTATGAGGAGCG 300
Db 241 TCGGGATGATACATGTCAGAACGACTGCTCAACTCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCAACCCCGGTCGTCCTGCTCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCAACCCCGGTCGTCCTGCTCGGAGAACAACTCTTCCCGC 360
QY 361 TGTGGTGGTACGCTCAACCCCGGTCGTCAGCTAGGAGAACCCAGCGTCCCCACGACA 420
Db 361 TGTGGTGGTACGCTCAACCCCGGTCGTCAGCTAGGAGAACCCAGCGTCCCCACGACA 420
QY 421 ATACGAGCGCAGCTCGAATTCACAGCTGTTCACCATCTCGCTCGCGCATGAGCGGTG 480
|||||
```

```
Db 421 ATACGAGCCACGTCGATTCGCCAGCTGTTCAACCATCTCCGCTCGCGGCATGAGCGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATATCCGCGCCACATAACGGGTGTCACGCTATGGCTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATATCCGCGCCACATAACGGGTGTCACGCTATGGCTGGGAT 540
QY 541 ATGATGATGAAGTGGT 556
|||||
Db 541 ATGATGATGAAGTGGT 556
|||||
RESULT 8
AX685026
LOCUS AX685026 606 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 25 from Patent WO0205548.
ACCESSION AX685026
VERSION AX685026.1 GI:29371431
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Maertens,G., Bosman,F. and Buyse,M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 25 18-JUL-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..606
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..603
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD86527.1"
/db_xref="GI:29371432"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLDGG
VNYATGNLPGCSFSIFLLALLSCLTVPASAYEVNVMYHVTNDCSNSIVVEAADM
INHTGCVPCVRENSSRCWVALPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIVPGHITGHRMAMWMMNWSPTTALVVSQLLRL"
mat_peptide 1..600
BASE COUNT 109 a 193 c 167 g 137 t
ORIGIN
Query Match 99.1%; Score 556; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTGCATCATACCCCTTACATCGGCTTCGCCACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAGTGCATCATACCCCTTACATCGGCTTCGCCACCTCGTGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTTGGCGCATGGCTCCGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTTGGCGCATGGCTCCGGTT 120
QY 121 CTGGAGGAGCGGCTGAACCTATGCAACAGGGAATTTGCCGGTGTCTTTCTATCTTC 180
Db 121 CTGGAGGAGCGGCTGAACCTATGCAACAGGGAATTTGCCGGTGTCTTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
Db 181 CTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
QY 241 TCGGGATGATACATGTCAGAACGACTGCTCAACTCAAGCATTTGTATGAGGAGCG 300
Db 241 TCGGGATGATACATGTCAGAACGACTGCTCAACTCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCAACCCCGGTCGTCCTGCTCGGAGAACAACTCTTCCCGC 360
|||||
```


b		CTCTTGGCTTTGCTGTCTCTGTCTGACCGTTCAGCCTTCGCCTTATGAAGTGCCCAACGTG	240
y		TCCGGGATGATACCATTGTCACGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGGCAGCG	300
b		TCCGGGATGTACCATGTTCACGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGGCAGCG	300
y		GACATGATCATGCACACCCGCCGGTGCCTGCTGCGGTTCGGGAGAACAACTCTTCCCCG	360
b		GACATGATCATGCACACCCGCCGGTGCCTGCTGCGGTTCGGGAGAACAACTCTTCCCCG	360
y		TGCTGGGTAGCGCTCACCCCACGCTCGCAGCTTAGNAAGCCACAGCGTCCCCACACGACA	420
b		TGCTGGGTAGCGCTCACCCCACGCTCGCAGCTTAGNAAGCCACAGCGTCCCCACACGACA	420
y		ATACGAGCGCCACGTCGATTCACCATCTTCACCATCTCGCCTCGCGCGCATGACAGCGTG	480
b		ATACGAGCGCCACGTCGATTCACCATCTTCACCATCTCGCCTCGCGCGCATGACAGCGTG	480
y		CAGGACTGCAATTGCTCAAATCTATCCCGGCACATAACGGGTCAACCGTATGGCTTGGGAT	540
b		CAGGACTGCAATTGCTCAAATCTATCCCGGCACATAACGGGTCAACCGTATGGCTTGGGAT	540
y		ATGATGATGAACCTGGT	556
b		ATGATGATGAACCTGGT	556
Db			
 RESULT 10			
AR157339			
LOCUS			
DEFINITION			
Sequence 27 from patent US 6245503.			
ACCESSION			
AR157339.1 GI:16218273			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Unclassified.			
REFERENCE			
AUTHORS			
TITLE			
Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use			
JOURNAL			
Patent: US 6245503-A 27 12-JUN-2001;			
FEATURES			
source			
1..636			
BASE COUNT			
119 a 203 c 174 g 140 t			
ORIGIN			
Query Match			
Best Local Similarity 99.1%; Score 556; DB 6; Length 636;			
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGTTGGTAAAGTCAATCGATACCCCTTACATCGGCTTCGCCAACCCTCGTGGGGTACATT	60
Db	1	ATGTTGGTAAAGTCAATCGATACCCCTTACATCGGCTTCGCCAACCCTCGTGGGGTACATT	60
Qy	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCAGGGCCCTGGCGCATGCGTCCGGGT	120
Db	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCAGGGCCCTGGCGCATGCGTCCGGGT	120
Qy	121	CTGGAGACGGGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTC	180
Db	121	CTGGAGACGGGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTC	180
Qy	181	CTCTTGGCTTTGCTGTCCTGTGTACCGTTCCAGTTCCCGTTATGAAGTCGCAACGTG	240
Db	181	CTCTTGGCTTTGCTGTCCTGTGTACCGTTCCAGTTCCCGTTATGAAGTCGCAACGTG	240
Qy	241	TCCGGGATGACATGTGTCAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGCAGCG	300
Db	241	TCCGGGATGACATGTGTCAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGCAGCG	300


```
QY 301 GACATGATCATGCACACCCCGGGTGGTGGTGGTGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGGTGGTGGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGTAGCGCTCACCCCGCCAGCTCGAGCTAGGAAGCCAGCGTCCCCACACGACA 420
Db 361 TCGTGGTAGCGCTCACCCCGCCAGCTCGAGCTAGGAAGCCAGCGTCCCCACACGACA 420
QY 421 ATACGAGCCAGCTCGATTCACCGCTGTTCAACATCTCGCTCGCGGAGATGAGCGGTG 480
Db 421 ATACGAGCCAGCTCGATTCACCGCTGTTCAACATCTCGCTCGCGGAGATGAGCGGTG 480
QY 481 CAGGACTGCAATTCCTCAATCTATCCCGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTCCTCAATCTATCCCGCCACATAACGGGTACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556

RESULT 11
AX452776
LOCUS 636 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 27 from Patent EP1211315.
ACCESSION AX452776
VERSION AX452776.1 GI:21712461
KEYWORDS Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 27 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..636
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..633
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38075.1"
/db_xref="GI:21712462"
VNYATGNLPGCSFSLFLLALLSCLTVPASAYEVNVMYHVNDCNSSIIVYEADM
IMHTPGCPVCEVRENSRCWALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIYPGHITGHRMADMMNWSPTTALVVSQLLRIVIEGRHHHHH
mat_peptide
1..630
/product="unnamed"
BASE COUNT 119 a 203 c 174 g 140 t
ORIGIN
Query Match 99.1%; Score 556; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATCGCGTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATCGCGTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGCGGCCCTAGGGGCGCTGCCAGGCGCTGCGCATGGCGTCCGGTT 120
Db 61 CCGCTCGTGGCGGCCCTAGGGGCGCTGCCAGGCGCTGCGCATGGCGTCCGGTT 120
QY 121 CTGGAGGACGGGTGAACCTATGCAACAGGGAATTTGCCGGTGTGCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGGTGAACCTATGCAACAGGGAATTTGCCGGTGTGCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCCTGTGACCGTTTCAGCTTCCGCTTATGAAGTGGCAACGTG 240
```

```
Db 181 CTCTTGGCTTTGCTGTCCTGTGTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
QY 241 TCGGAGATGTACCATGTACGAAACGACTGCTCCAATCTCAAGCATTTGTGTATGAGGAGCG 300
Db 241 TCGGAGATGTACCATGTACGAAACGACTGCTCCAATCTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGTGGTGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGGTGGTGGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGTAGCGCTCACCCCGCCAGCTCGAGCTAGGAAGCCAGCGTCCCCACACGACA 420
Db 361 TCGTGGTAGCGCTCACCCCGCCAGCTCGAGCTAGGAAGCCAGCGTCCCCACACGACA 420
QY 421 ATACGAGCCAGCTCGATTCACCGCTGTTCAACATCTCGCTCGCGGAGATGAGCGGTG 480
Db 421 ATACGAGCCAGCTCGATTCACCGCTGTTCAACATCTCGCTCGCGGAGATGAGCGGTG 480
QY 481 CAGGACTGCAATTCCTCAATCTATCCCGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTCCTCAATCTATCCCGCCACATAACGGGTACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556

RESULT 12
AX685028
LOCUS 636 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 27 from Patent WO0205548.
ACCESSION AX685028
VERSION AX685028.1 GI:29371433
KEYWORDS Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F. and Buyse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 27 18-JUL-2002;
Innogenetics N.V. (BE)
FEATURES
Location/Qualifiers
source
1..636
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..633
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD86528.1"
/db_xref="GI:29371434"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLVDG
VNYATGNLPGCSFSLFLLALLSCLTVPASAYEVNVMYHVNDCNSSIIVYEADM
IMHTPGCPVCEVRENSRCWALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIYPGHITGHRMADMMNWSPTTALVVSQLLRIVIEGRHHHHH"
mat_peptide
1..630
/product="unnamed"
BASE COUNT 119 a 203 c 174 g 140 t
ORIGIN
Query Match 99.1%; Score 556; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATCGCGTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATCGCGTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGCGGCCCTAGGGGCGCTGCCAGGCGCTGCGCATGGCGTCCGGTT 120
```



```
Db 1 ATGTTGGGTAAGGTACATCGATACCCCTTACATCGGGCTTCGCCGACCTCGTGGGTACATT 60
Qy 61 CCGCTCGTCGGCGCCCCCTAGGGGCGGTGTCAGGGGCCCTGCGCATGCGTCGCGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGCGGTGTCAGGGGCCCTGCGCATGCGTCGCGGTT 120
Qy 121 CTGGAGGCGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTTCTCTATCTTC 180
Db 121 CTGGAGGCGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTTCTCTATCTTC 180
Qy 181 CTCTGGCTTTGCTGTCCTGCTGACCGTTCCAGCTTTCAGCTTTCAGCTTTCAGCTTTC 240
Db 181 CTCTGGCTTTGCTGTCCTGCTGACCGTTCCAGCTTTCAGCTTTCAGCTTTCAGCTTTC 240
Qy 241 TCCGGATGTACCATGTGACGAACAGTCTGCTCAACTCAAGCAATTTGTATGAGGAGCG 300
Db 241 TCCGGATGTACCATGTGACGAACAGTCTGCTCAACTCAAGCAATTTGTATGAGGAGCG 300
Qy 301 GACATGATCATGACACACCCCGGGTGCCTGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGCCTGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
Qy 361 TCGTGGTAGCGCTACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCGACACGACA 420
Db 361 TCGTGGTAGCGCTACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCGACACGACA 420
Qy 421 ATACGAGCGGACGTCGATTCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 ATACGAGCGGACGTCGATTCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTTCACCGTATGGCTTGG 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTTCACCGTATGGCTTGG 540
Qy 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556

RESULT 15
AX452770
LOCUS AX452770 723 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 21 from Patent EP1211315.
ACCESSION AX452770
VERSION AX452770.1 GI:21712455
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 21 05-JUN-2002;
IMMUNOGENETICS N.V. (BE)
FEATURES
source
1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38072.1"
/db_xref="GI:21712456"
/translation="MLGKVIDITLTCGFADLVGYIPLVGAPLGAARALAHGVRVLEEDG
VYATGNLPGCSFSLFLLALLSCLTPASAYEVNRYVSGMYHVNDCSNSSIVYEADM
IMHTPGCPVCRNNSRCWALPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
ODNCGSIYPGHITGHRMAMWMMNWSPTTALVVSQLLRIPQAVVDVAVGAHGVLAGL
AYYSVMGNWAKVLIVMLFAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
```

ORIGIN

```
Query Match 99.1%; Score 556; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGGGTAAGGTACATCGATACCCCTTACATCGGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAGGTACATCGATACCCCTTACATCGGGCTTCGCCGACCTCGTGGGTACATT 60
Qy 61 CCGCTCGTCGGCGCCCCCTAGGGGCGGTGTCAGGGGCCCTGCGCATGCGTCGCGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGCGGTGTCAGGGGCCCTGCGCATGCGTCGCGGTT 120
Qy 121 CTGGAGGCGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTTCTCTATCTTC 180
Db 121 CTGGAGGCGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTTCTCTATCTTC 180
Qy 181 CTCTGGCTTTGCTGTCCTGCTGACCGTTTCAGCTTTCAGCTTTCAGCTTTCAGCTTTC 240
Db 181 CTCTGGCTTTGCTGTCCTGCTGACCGTTTCAGCTTTCAGCTTTCAGCTTTCAGCTTTC 240
Qy 241 TCCGGATGTACCATGTGACGAACAGTCTGCTCAACTCAAGCAATTTGTATGAGGAGCG 300
Db 241 TCCGGATGTACCATGTGACGAACAGTCTGCTCAACTCAAGCAATTTGTATGAGGAGCG 300
Qy 301 GACATGATCATGACACACCCCGGGTGCCTGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGCCTGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
Qy 361 TCGTGGTAGCGCTACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCGACACGACA 420
Db 361 TCGTGGTAGCGCTACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCGACACGACA 420
Qy 421 ATACGAGCGGACGTCGATTCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 ATACGAGCGGACGTCGATTCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTTCACCGTATGGCTTGG 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTTCACCGTATGGCTTGG 540
Qy 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556
```

Search completed: December 20, 2003, 02:02:02
Job time : 2126.1 Secs

THIS PAGE BLANK (USPTO)

XX PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

PF 11-JAN-2002; 2002WO-EP00219.
XX 11-JAN-2001; 2001US-260699P.
PR 30-AUG-2001; 2001US-315768P.
XX (INNO-) INNOGENETICS NV.
XX Maertens G, Bosman F, Buyse M;
XX WPI; 2002-599657/64.
DR P-PSDB; AAO18668.
XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection
XX Example 2; Page 177-178; 243pp; English.
XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
PS Sequence 606 BP; 109 A; 193 C; 167 G; 137 T; 0 other;
SQ

Query Match 99.1%; Score 556; DB 24; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.2e-138;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGCGCCCTAGGGGGCTGCCAGGCGCTTGGCGCATGCGTCCGGTT 120
DB 61 CCGCTCGTGGCGCCCTAGGGGGCTGCCAGGCGCTTGGCGCATGCGTCCGGTT 120

QY 121 CTGAGGACGCGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGAGGACGCGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

QY 181 CTCCTTGGCTTTGCTGTCTGTACCGTTCCAGTTCGGTTTATGAAGTGGCAACGTG 240
DB 181 CTCCTTGGCTTTGCTGTCTGTACCGTTCCAGTTCGGTTTATGAAGTGGCAACGTG 240

QY 241 TCCGGGTATACCATGTACGACGACGTCTCAACTCAAGCATTTGTATGAGCGAGCG 300
DB 241 TCCGGGTATACCATGTACGACGACGTCTCAACTCAAGCATTTGTATGAGCGAGCG 300

QY 301 GACATGATCATGCACACCCCGGGTGCCTGCCCTGCTCGGTTCGGGAGAAACAATCTTCCGC 360
DB 301 GACATGATCATGCACACCCCGGGTGCCTGCCCTGCTCGGTTCGGGAGAAACAATCTTCCGC 360

QY 361 TCGTGGGTAGCGCTCACCCCGACGCTCGAGTAGGACGCGCGTCCCGACACGACA 420
DB 361 TCGTGGGTAGCGCTCACCCCGACGCTCGAGTAGGACGCGCGTCCCGACACGACA 420

QY 421 ATACGACGCGCATGTGATTCACAGCTGTTTCCATCTCGCTTCGCGGCATGAGCGGTG 480
DB 421 ATACGACGCGCATGTGATTCACAGCTGTTTCCATCTCGCTTCGCGGCATGAGCGGTG 480

QY 481 CAGGACTGCAATGTCTCAATATATCCCGGCGCATTAACGGGTACCGTATGGCTTGGAT 540
DB 481 CAGGACTGCAATGTCTCAATATATCCCGGCGCATTAACGGGTACCGTATGGCTTGGAT 540

QY 541 ATGATGATGAACCTGGT 556
DB 541 ATGATGATGAACCTGGT 556

RESULT 5
AAT12964
ID AAT12964 standard; DNA; 636 BP.
XX
AC AAT12964;
XX 24-SEP-1996 (first entry)
XX HCV E1 construct HCC140.
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX Hepatitis C virus.
XX WO9604385-A2.
XX 15-FEB-1996.
XX 31-JUL-1995; 95WO-EP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
DR
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di: sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.

AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
and E2 protein coding sequence constructs. These sequences are included
in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
The recombinant proteins can then be isolated using a method of the
invention. In the method, the envelope proteins are purified by
carrying out a disulphide bond cleavage, or a reduction step with a
disulphide bond cleavage agent, after lysis of recombinant host cells.
The constructs containing the purified HCV envelope proteins can be used
for vaccinating humans against HCV, for in vitro detection of HCV
antibodies in a sample, and in a serotyping assay for detecting one or
more serological types of HCV present in a biological sample. The
constructs can also be immobilised on a solid substrate and incorporated
into a reversed phase hybridisation assay for determining the presence or
the genotype of HCV. The new purification method preserves the
conformation of the recombinantly expressed E1, E2 and E1/E2, and
eliminates contaminating proteins. Antigens isolated using this method
are more reactive with human sera than those isolated by known
techniques.

Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;

Query Match 99.1%; Score 556; DB 17; Length 636;
Best Local Similarity 100.0%; Pred. No. 1.2e-138;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGCGCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGTCCGGTT 120

Query Match 99.1%; Score 556; DB 24; Length 723;
Best Local Similarity 100.0%; Pred. No. 1.3e-138;

601 ATGGCTGGGATATGATGATGAACCTGGT 628

XX	AA148914;	
AC		
XX		
DT	24-OCT-2002	(first entry)
XX		
DE	Hepatitis C virus clone HCC110A E1 protein coding sequence.	
XX		
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene; immunosuppressant; vaccine; ds.	

XX	Hepatitis C virus.
OS	
XX	
XX	WO200255548-A2.
PN	
XX	
XX	18-JUL-2002.
PD	

11-JAN-2002; 2002WO-EP00219.
11-JAN-2001; 2001US-260699P.
30-AUG-2001; 2001US-315768P.

XX
PA (INNO-) INNOGENETICS NV.
XX
PI Maertens G. Bosman F, Buyse M;

XX
DR WPI; 2002-599657/64.
DR P-PSDB: AAO18661.

DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.

XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;

PA (INNO-) INNOGENETICS NV.
XX
XX Maartens G. Bosman F. Buysse M;
XX

XX
ne
wpt : 2002-599657/64.

DR P-PSDE; AA018678.

PT New therapeutic vaccine compositions comprising at least one specific oligomeric recombinant hepatitis C virus (HCV) single or specific for immunizing humans

PT	recombinant envelope protein E1 of 22, about 100 kDa
PT	from HCV infection

xx Example 2; Page 206-209; 243pp; English.

CC The present invention relates to new therapeutic vaccine compositions
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a

CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ

Query Match 84.0%; Score 471; DB 24; Length 2082;
Best Local Similarity 88.5%; Pred. No. 8.4e-116;
Matches 553; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 4 TTGGGTAAGGTCTATCGATACCCCTTACATGCGGGTTCCGCCGACCTCGTGGGTACATTCCG 63
Db 4 TTGGGTAAGGTCTATCGATACCCCTTACATGCGGGTTCCGCCGACCTCGTGGGTACATTCCG 63
QY 64 CTGCTGCGGCCGCCCTTAGGGGGCGGTGCCAGGGCCCTGCGCATGCGTCCGGGTCTG 123
Db 64 CTGCTGCGGCCGCCCTTAGGGGGCGGTGCCAGGGCCCTGCGCATGCGTCCGGGTCTG 123
QY 124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTCCCGGGTTCCTTCTATCTTCCTC 183
Db 124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTCCCGGGTTCCTTCTATCTTCCTC 183
QY 184 TTGGCTTTTGTCTGCTGTCTGACCGTTCCAGCTTCGCTTATGAAGTGCACAGTGTCC 243
Db 184 TTGGCTTTTGTCTGCTGTCTGACCGTTCCAGCTTCGCTTATGAAGTGCACAGTGTCC 243
QY 244 GGGATGTACATGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
Db 244 GGGATGTACATGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
QY 304 ATGATCATGCACACCCCGGGTTCGCTGCTGTTCCGGAGAACAACTTCTCCCGTGC 363
Db 304 ATGATCATGCACACCCCGGGTTCGCTGCTGTTCCGGAGAACAACTTCTCCCGTGC 363
QY 364 TGGGTAGCGCTCACCCGCCGCTCGAGCTAGGAACGCCAGCGTCCCCACACGACAATA 423
Db 364 TGGGTAGCGCTCACCCGCCGCTCGAGCTAGGAACGCCAGCGTCCCCACACGACAATA 423
QY 424 CGACGCCACGTCGAT----- 438
Db 424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTCTTCTGTTCCGCTATGATGTTGGG 483
QY 439 -----TCCACGCTGTTCAACCATCTCGCTCGCCGCGCAT 471
Db 484 GACCTCTGCGGATCTGTCTTCTCTGCTCCAGCTGTTCAACCATCTCGCTCGCCGCGCAT 543
QY 472 GAGACGCTGCAGACTGCAATTGCTCAATCTATCCGGCCACATACGGGTACCCGTATG 531
Db 544 GAGACGCTGCAGACTGCAATTGCTCAATCTATCCGGCCACATACGGGTACCCGTATG 603
QY 532 GCTTGGGATATGATGATGAACTGGT 556
Db 604 GCTTGGGATATGATGATGAACTGGT 628

RESULT 12

AA12973
ID AA12973 standard; DNA; 2086 BP.
XX
AC AA12973;
XX
DT 24-SBP-1996 (first entry)
XX
DE HCV E1 construct HCC165.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.

Hepatitis C virus.

WO9604385-A2.

15-FEB-1996.

31-JUL-1995; 95WO-EP03031.

29-JUL-1994; 94EP-0870132.

(INNO-) INNOGENETICS NV.

Bosman P, Buyse M, De Martynoff G, Maertens G;

WPI; 1996-129401/13.

Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

proteins - in presence of di: sulphide bond cleavage agent, to

produce proteins suitable for direct use in vaccines or diagnostic

assays of HCV

Claim 23; Fig 21; 146pp; English.

AA12704-T12709 and AA12961-T12974 represent hepatitis C virus (HCV) E1

and E2 protein coding sequence constructs. These sequences are included

in vectors for the production of recombinant E1, E2, and E1/E2 proteins.

The recombinant proteins can then be isolated using a method of the

invention. In the method, the envelope proteins are purified by

carrying out a disulphide bond cleavage, or a reduction step with a

disulphide bond cleavage agent, after lysis of recombinant host cells.

The constructs containing the purified HCV envelope proteins can be used

for vaccinating humans against HCV, for in vitro detection of HCV

antibodies in a sample, and in a serotyping assay for detecting one or

more serological types of HCV present in a biological sample. The

constructs can also be immobilised on a solid substrate and incorporated

into a reversed phase hybridisation assay for determining the presence or

the genotype of HCV. The new purification method preserves the

conformation of the recombinantly expressed E1, E2 and E1/E2, and

eliminates contaminating proteins. Antigens isolated using this method

are more reactive with human sera than those isolated by known

techniques.

XX

SQ

Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;

Query Match 84.0%; Score 471; DB 17; Length 2086;

Best Local Similarity 88.5%; Pred. No. 8.4e-116;

Matches 553; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 4 TTGGGTAAGGTCTATCGATACCCCTTACATGCGGGTTCCGCCGACCTCGTGGGTACATTCCG 63

Db 4 TTGGGTAAGGTCTATCGATACCCCTTACATGCGGGTTCCGCCGACCTCGTGGGTACATTCCG 63

QY 64 CTGCTGCGGCCGCCCTTAGGGGGCGGTGCCAGGGCCCTGCGCATGCGTCCGGGTCTG 123

Db 64 CTGCTGCGGCCGCCCTTAGGGGGCGGTGCCAGGGCCCTGCGCATGCGTCCGGGTCTG 123

QY 124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTCCCGGGTTCCTTCTATCTTCCTC 183

Db 124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTCCCGGGTTCCTTCTATCTTCCTC 183

QY 184 TTGGCTTTTGTCTGCTGTCTGACCGTTCCAGCTTCGCTTATGAAGTGCACAGTGTCC 243

Db 184 TTGGCTTTTGTCTGCTGTCTGACCGTTCCAGCTTCGCTTATGAAGTGCACAGTGTCC 243

QY 244 GGGATGTACATGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303

Db 244 GGGATGTACATGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303

QY 304 ATGATCATGCACACCCCGGGTTCGCTGCTGTTCCGGAGAACAACTTCTCCCGTGC 363

Db 304 ATGATCATGCACACCCCGGGTTCGCTGCTGTTCCGGAGAACAACTTCTCCCGTGC 363

the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;

Query Match 84.0%; Score 471; DB 17; Length 2433;
Best Local Similarity 88.5%; Pred. No. 8.8e-116; Indels 72; Gaps 1;
Matches 553; Conservative 0; Mismatches 0;

4 TTGGGTAAAGTCAATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCG 63
355 TTGGGTAAAGTCAATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCG 414
64 CTCGTGCGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123
415 CTCGTGCGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 474
124 GAGACGGCGTGAACATATGCAACAGGGAATTTGGCCCGTTGCTCTTCTCTATCTTCTC 183
475 GAGACGGCGTGAACATATGCAACAGGGAATTTGGCCCGTTGCTCTTCTCTATCTTCTC 534
184 TTGGCTTTGCTGCTCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTGTC 243
535 TTGGCTTTGCTGCTCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTGTC 594
244 GGGATGTACCATGTGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 303
595 GGGATGTACCATGTGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 654
304 ATGATCATGCACACCCCGGGTGGCTGCGCTTCCGGGAGAACAACTTTCCCGCTGTC 363
655 ATGATCATGCACACCCCGGGTGGCTGCGCTTCCGGGAGAACAACTTTCCCGCTGTC 714
364 TGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGCGTCCCAACCAACAATA 423
715 TGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGCGTCCCAACCAACAATA 774
424 CGACGCCACGTGCTGATTTGCTTGGGGCGGCTGCTTTCTGCTATGATGCTGGGG 834
775 CGACGCCACGTGCTGATTTGCTTGGGGCGGCTGCTTTCTGCTATGATGCTGGGG 894
439 -----TCCAGCTGTTTACCATCTCGCTCGCGGCAAT 471
835 GACCTCTGCGGATCTGCTTCTTCTGCTCCAGCTGTTTACCATCTCGCGTCCGGGCAAT 894
472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 531
895 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 954
532 GCTTGGGATATGATGATGAACTGCT 556
955 GCTTGGGATATGATGATGAACTGCT 979

RESULT 14
AAT12706
ID AAT12706 standard; DNA; 2433 BP.
AC AAT12706;
XX AAT12706;
XX 23-SEP-1996 (first entry)
XX HCV E1 construct HCCI11A.
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX ss.
XX Hepatitis C virus.

304 ATGATCATGCACACCCCGGGTGGCTGCGCTTCCGGGAGAACAACTTTCCCGGTCG 363
364 TGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGCGTCCCAACCAACAATA 423
364 TGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGCGTCCCAACCAACAATA 423
424 CGACGCCACGTGCTGATTTGCTTGGGGCGGCTGCTTTCTGCTATGATGCTGGGG 483
424 CGACGCCACGTGCTGATTTGCTTGGGGCGGCTGCTTTCTGCTATGATGCTGGGG 483
439 -----TCCAGCTGTTTACCATCTCGCTCGCGGCAAT 471
484 GACCTCTGCGGATCTGCTTCTTCTCCAGCTGTTTACCATCTCGCGTCCGGGCAAT 543
472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 531
544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 603
532 GCTTGGGATATGATGATGAACTGCT 556
604 GCTTGGGATATGATGATGAACTGCT 628

RESULT 13
AAT12974
ID AAT12974 standard; DNA; 2433 BP.
AC AAT12974;
XX AAT12974;
XX 25-SEP-1996 (first entry)
XX HCV E1 construct HCCI166.
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX ss.
XX Hepatitis C virus.
XX WO9604385-A2.
XX 15-FEB-1996.
XX 31-JUL-1995; 95WO-EP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
XX proteins - in presence of disulphide bond cleavage agent, to
XX produce proteins suitable for direct use in vaccines or diagnostic
XX assays of HCV
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or

Mon Dec 22 13:28:39 2003

Matches 550;		Conservative 0;	Mismatches 11;	Indels 72;	Gaps 1;
Qy	1	ATGTTGGGTAAGTATCATGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT	60		
Db	1	ATGTTGGGTAAGTATCATGATACCCCTTACATCGGCTTCGCCGACCTCATGGGGTACATT	60		
Qy	61	CCGCTCGTCGGCGCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT	120		
Db	61	CCGCTCGTCGGCGCCCTAGGGGGGTGTGCAGAGCCCTGGCGCATGGCGTCCGGGTT	120		
Qy	121	CTGGAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC	180		
Db	121	CTGGAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC	180		
Qy	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCGAACGCTG	240		
Db	181	CTCTTGGCTTTTACTGTCTCTGTGACCAATTCAGCTTCCGCTTATGAAGTGGCGAACGCTG	240		
Qy	241	TCCGGGATGTACCATGTACGAAACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGG	300		
Db	241	TCCGGGATGTACCATGTACGAAACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGG	300		
Qy	301	GACATGATCATGACACACCCCGGGTGCCTGCGTTCCGGAGAACACTCTTCCCGC	360		
Db	301	GACATGATCATGACACACCCCGGGTGCCTGCGTTCCGGAGAACACTCTTCCCGC	360		
Qy	361	TGCTGGGTAGCGCTCACCCCGCTCGAGCTAGGAACGCCAGCGTCCCGCACGACA	420		
Db	361	TGCTGGGTAGCGCTCACCCCGCTCGAGCTAGGAACGCCAGCGTCCCGCACGACA	420		
Qy	421	ATACGACGCCACGTGAT-----	438		
Db	421	ATACGACGCCACGTGATTTGCTGTTGGGGCGCTGCTTTCTGTTCCGCTATGTACGTG	480		
Qy	439	-----TCCGAGCTGTTACCATCTCGCCTCGCGG	468		
Db	481	GGGGATCTCTGGGATCTGTCTTCCTGCTCTCCAGCTGTTCAACATCTCGCCTCGCGG	540		
Qy	469	CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGT	528		
Db	541	CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGT	600		
Qy	529	ATGGCTTGGGATATGATGATGAACCTGTAATAG	561		
Db	601	ATGGCTTGGGATATGATGATGAACCTGTAATAG	633		

Search completed: December 19, 2003, 18:51:15
Job time : 158.093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1443.21 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-23
Perfect score: 561
Sequence: 1 ATGTCGGTAAGTCAATCGA.....TGATGTAAGTGGTAATAG 561

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	62.6	11.2	488	9 AV755731	AV755731 AV755731
C 2	50.6	9.0	432	9 AV758366	AV758366 AV758366
C 3	41.6	7.4	502	12, BI879124	BI879124 fm04608.y
C 4	40.6	7.2	275	9 AV835132	AV835132 AV835132

5	40.6	7.2	402	9	AV392783	AV392783 AV392783
6	40.6	7.2	551	9	AV392165	AV392165 AV392165
7	40.6	7.2	552	12	BI996341	BI996341 1031037A0
8	40.6	7.2	584	12	BI727879	BI727879 1031035C1
9	40.4	7.2	1201	13	BI356664	BI356664 BX356664
10	40.2	7.2	1162	12	BM918259	BM918259 AGENCOURT
11	40	7.1	1201	9	AL513886	AL513886 AL513886
12	39	7.0	359	12	BJ252869	BJ252869 BJ252869
13	39	7.0	375	12	BJ246716	BJ246716 BJ246716
14	39	7.0	840	29	CC335916	CC335916 OGUAJ60TV
15	39	7.0	873	14	CD446071	CD446071 EL01T0207
16	38.6	6.9	925	29	CNS0091P	AL053013 Drosophila
17	38.4	6.8	636	12	BI960110	BI960110 HVSME002
18	38.4	6.8	702	14	CD432549	CD432549 ETH1_30_D
19	38.4	6.8	970	29	CNS010C9	AL098787 Drosophila
20	38.4	6.8	987	29	CNS015VX	AL105975 Drosophila
21	38.2	6.8	533	6	AU192776	AU192776 Porphyra
22	38.2	6.8	538	6	AU193705	AU193705 Porphyra
23	38.2	6.8	544	6	AU190971	AU190971 Porphyra
24	38.2	6.8	1270	12	BG968359	AU192419 Porphyra
25	38.2	6.8	1201	13	BX381961	BX381961 BX381961
26	38	6.8	1201	13	BX381961	BX381961 BX381961
27	38	6.7	435	14	C72860	C72860 C72860 Rice
28	37.8	6.7	533	29	CC010084	CC010084 FUGJ921B
29	37.8	6.7	659	29	CC405164	CC405164 FUGJ921B
30	37.8	6.7	826	29	BZ736582	BZ736582 OGER42TC
31	37.8	6.7	895	29	CC359028	CC359028 PUHFD18TD
32	37.8	6.7	925	29	CC359026	CC359026 PUHFD18TD
33	37.8	6.7	940	29	CC010085	CC010085 FUGJ921D
34	37.8	6.7	951	29	CC405167	CC405167 FUGJ921D
35	37.6	6.7	431	9	AV639153	AV639153 AV639153
36	37.6	6.7	360	9	AJ473805	AJ473805 AJ473805
37	37.4	6.7	637	13	BQ293470	BQ293470 1091016F0
38	37.4	6.7	641	13	BQ172543	BQ172543 1091024A1
39	37.4	6.7	650	14	CA828039	CA828039 1114022H0
40	37.4	6.7	834	29	BZ641450	BZ641450 OGCCS41TC
41	37.4	6.7	841	29	BZ641457	BZ641457 OGCCS41TC
42	37.4	6.7	856	29	BZ578381	BZ578381 msh2 5817
43	37.4	6.7	872	29	BZ555011	BZ555011 pacal-60
44	37.4	6.7	563	10	BE490055	BE490055 WHE0364_A
45	37.2	6.6				

ALIGNMENTS

RESULT 1	AV755731/c	488 bp	linear	EST 19-OCT-2000
LOCUS	AV755731	BM Homo sapiens	CDNA clone	BMFAKB03 5', mRNA sequence.
DEFINITION	AV755731			
ACCESSION	AV755731			
VERSION	AV755731.1	GI:10913579		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 488)			
AUTHORS	Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.			
TITLE	Homo sapiens CDNA BM clones			
JOURNAL	Unpublished			
COMMENT	Contact: Zeguaguan Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex. 45) Fax: 86-21-50801922 Email: han@chgc.sh.cn This clone is available at CHGC in Shanghai. Location/Qualifiers			


```

source
1. .488
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN
Query Match 11.2%; Score 62.6; DB 9; Length 488;
Best Local Similarity 77.9%; Pred. No. 1e-05;
Matches 88; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 445 CTGTTACCATCTCGCTCCCGGATGAGACGGTGCAGGCTGCAATGCTCAATCTAT 504
Db 403 CAGCTGATCATCTGGCTCAGACCATGAGTTGTGCATGATGCAACTGCTCCATCTAT 344
QY 505 CCGGCGCACATAACGGTGTACCGTATG-GCTTGGGATATGATGCAACTGGT 556
Db 343 CCTGGCCCATCACTGACACCGTATGAGCATGGGACATGATGCACTGGT 291

RESULT 2
AV758366/c 492 bp mrna linear EST 19-OCT-2000
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAKA03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 492)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
TITLE Homo sapiens cDNA BM clones
JOURNAL Unpublished
COMMENT Contact: Zeguang Han
Chinese National Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1. .492
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="BMFAKA03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN
Query Match 9.0%; Score 50.6; DB 9; Length 492;
Best Local Similarity 72.5%; Pred. No. 0.013;
Matches 79; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

QY 449 TCACCATCTCGCTCCCGGATGAGACGGTGCAGGCTGCAATGCTCAATCTATCCCG 508
Db 400 TGATTATCTCTCAGCAGCAACATGGTTGTGCAAGATGCACTGCTATCTATCCTG 341

```

```

QY 509 GCCACATAACGGTC-ACCGTATGGCTTGGGATATGATGATGAACTGGT 556
Db 340 GCTGCATCACTGGACTACAGTATGGCATAGGCTATGATGATGAACTGGT 292

RESULT 3
BI879124/c 502 bp mrna linear EST 13-FEB-2002
LOCUS BI879124
DEFINITION fno4608.y1 Zebrafish adult retina cDNA Danio rerio cDNA clone
IMAGE:4145367 5', similar to TR:Q9PWN4 Q9PWN4 RHODOPSIN.; mRNA
sequence.
ACCESSION BI879124
VERSION BI879124.1 GI:16086395
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 502)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
JOURNAL Unpublished
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
Library constructed by: Chandra Tucker and Gregory Niemi DNA
Sequencing by: Washington University Genome Sequencing Center Clone
distribution: RessourcenzentrumPrimarDatenbank, Berlin, Germany
(web address: www.rzpd.de)
Trace considered overall poor quality
Seq primer: T3 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .502
/organism="Danio rerio"
/mol_type="mrna"
/strain="wild-type"
/db_xref="taxon:7955"
/clone="IMAGE:4145367"
/sex="mixed"
/dev_stage="1-2 years"
/lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')"
/clone_lib="Zebrafish adult retina cDNA"
/notes="Vector: Lambda ZAP II (phagescript SK-); Site_1:
ECORI; Site 2: SalI; This Zebrafish library was
constructed by Dr. Susan E. Brockerhoff (email:
sbrocker@u.washington.edu) RZPD library number: 760"
BASE COUNT 98 a 163 c 125 g 116 t
ORIGIN
Query Match 7.4%; Score 41.6; DB 12; Length 502;
Best Local Similarity 51.0%; Pred. No. 2.8;
Matches 98; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 300 GGACATGATATGCACACCCCGGTCGGTCCGTCGGAGAACCAACTCTTCCCG 359
Db 484 GGCCATGAGCATGTGAAGCCACCACATGATGGATGGTTTCGCCAGACCGAAAGTTGTTG 425
QY 360 CTGCTGGGTAGCGTCTACCCCGGCTCGCAGCTAGGAACGCCAGCGTCCCGACGAC 419
Db 424 CCAGTTTGCACTCCACCATCCAGCGCTCAATGGCATGTACTCGAGCGACCATAGCCCAT 365
QY 420 AATAGCGCCACCGTCGATGCCAGTTCCTCCAGCTTTCACCATCTCGCCTCGCGGATGAGACGGT 479

```



```

Db      364 CTACCGCGGAGGCTGCTAAGTAGCCTTACAGGTTGACGCTCAGTCGGCAGAACCGAA 305
QY      480 GCAGGACTGCAA 491
Db      304 GTAGCGGTGCAA 293

RESULT 4
AV835132
LOCUS   AV835132
DEFINITION AV835132 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
spontaneum top three leaves adult, heading stage Hordeum vulgare
subsp. spontaneum cDNA clone bah24018, mRNA sequence.
ACCESSION AV835132
VERSION   AV835132.1 GI:14527221
KEYWORDS  EST.
SOURCE    Hordeum vulgare subsp. spontaneum
ORGANISM  Hordeum vulgare subsp. spontaneum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
          ; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 275)
AUTHORS   Sato, K.
TITLE     Barley EST sequencing project in NIG and Okayama Univ
JOURNAL   Unpublished
COMMENT   Contact: Kazuhiro Sato
          Research Institute for Bioresources
          Okayama University, Barley Germplasm Center
          Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
          Email: kazuato@rib.okayama-u.ac.jp/
          URL: http://www.rib.okayama-u.ac.jp/barley/
          Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct
          submission;
          database: http://www.shigen.nig.ac.jp/barley/Barley.html.

FEATURES             source
    Location/Qualifiers
        1..275
            /organism="Hordeum vulgare subsp. spontaneum"
            /mol_type="mRNA"
            /cultivar="H602"
            /db_xref="taxon:77009"
            /clone="bah24018"
            /issue_type="top three leaves"
            /dev_stage="adult, heading stage"
            /clone_lib="K. Sato unpublished cDNA library: Hordeum
            vulgare subsp. spontaneum top three leaves adult, heading
            stage"
BASE COUNT  30 a 123 c 67 g 46 t 9 others
ORIGIN
Query Match      7.2%; Score 40.6; DB 9; Length 275;
Best Local Similarity 48.2%; Pred. No. 4.3;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY      306 GATCATGCACACCCCGGTGCGTCCGCTCGGTCGGGAGAACAACTTCCCGCTGCTG 365
Db      1  GGTCTCTCGACGGAACCCGCCCTCTGCTCCGTTCTGCTCCCTGCGTCCGACCGCCG 60

QY      366 GGTAGCGCTCACCCACGCTCGACGTAGGACGCGCGGTCCACACGACAAATACG 425
Db      61  CGGCCAGCGGCACTCACCTNCTCTGCTCGTTGCGCCCTCCCTCTCGCGCCCGCCCTCG 120

QY      426 AGCCACGCTCGATTCCTCAGCTGTTCACCATCTCGCTCGCGGCATGACGCGTGCAGGA 485
Db      121 ACGGCTGGCGTAGCTGGGTGAGTTACATCTCTCTCGGACGCGGACGCTGCGCGA 180

QY      486 CTGCAATTGCTCAATATATCCCGGCGCACATAACGGGTTCACCGTATG 531
Db      181 ACCGNACTCGCTCCCTCTCTCGGCCCGCAGGCTCGCACCGGAAACG 226

RESULT 5
AV392783
LOCUS   AV392783
DEFINITION AV392783 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CM096g04_r 5', mRNA sequence.
ACCESSION AV392783
VERSION   AV392783.1 GI:6546999
KEYWORDS  EST.
SOURCE    Chlamydomonas reinhardtii
ORGANISM  Chlamydomonas reinhardtii
          Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
          Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 402)
AUTHORS   Aamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
TITLE     A large scale structural analysis of cDNAs in a unicellular green
          alga, Chlamydomonas reinhardtii. I. Generation of 3433
          non-redundant expressed sequence tags
JOURNAL   DNA Res. 6 (6), 369-373 (1999)
MEDLINE   20152988
PUBMED    10691129
COMMENT   Contact: Yasukazu Nakamura
          The First Laboratory for Plant Gene Research
          Kazusa DNA Research Institute
          Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
          Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES             source
    Location/Qualifiers
        1..402
            /organism="Chlamydomonas reinhardtii"
            /mol_type="mRNA"
            /strain="C9"
            /db_xref="taxon:3055"
            /clone="CM096g04_r"
            /dev_stage="photoautotrophic growth"
            /clone_lib="Chlamydomonas reinhardtii C9"
            /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
            XhoI"
BASE COUNT  66 a 126 c 146 g 64 t
ORIGIN
Query Match      7.2%; Score 40.6; DB 9; Length 402;
Best Local Similarity 45.3%; Pred. No. 4.7;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY      42  CGACCTCTGGGGTACATTCCTCGCTCGTGGCGGCCCCCTAGGGGGGCGCTCCAGGGCCCT 101
Db      53  CGAGCTCATCTCTGTCTATTTGTGCGCGGCACTGAACATGAAGAGCGTGTGACGGACCT 112
QY      102 GCGCATGCGCTCCGGGTTCTGAGGAGCGCGTGAACATGCAACAGGGAATTTGCCCGG 161
Db      113 GCGCGCGCGCGCGGAGTGGGAGGGCGGCTACGGCGACGAGTCCGTGAGCTTGGGCGC 172
QY      162 TTGCTCTTTCTCTATCTTCTCTTGGCTTTGCTGTCTGTGACCGTTCAGCTTCCGCGC 221
Db      173 CGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGMACTCAAGGCCCGAAGACCCAG 232
QY      222 TTATGAGTGGCGAAGCTGTCGGGATGTACCATGTACAGAACGACTGCTTCAACTCAAG 281
Db      233 CTTTCGCGCTCGCTGCTGGGCGCACCTCGTGGCGGCGGCGACCGCGGCTGCTGCGAT 292
QY      282 CATTGTGTATGAGCGAGCGACATGATCATGACACCGCGGCTGCGTCCGCTCGCTCG 341
Db      293 CTTGATGACCGACGAGGAGTTTGGGCGCGCATCTACGGCGGCGGTGCCCATGCGCGG 352
QY      342 GGAGAACAACTTCTCCCGCTGCTGGGT 368
Db      353 CAAGAGACGAGGCGAGCTACATGAT 379

RESULT 6
AV392165
LOCUS   AV392165
DEFINITION AV392165 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CM083e05_r 5', mRNA sequence.
ACCESSION AV392165
VERSION   AV392165.1 GI:6546381

```


KEYWORDS
 EST.
 Chlamydomonas reinhardtii
 Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadaceae; Chlamydomonas.
REFERENCE
AUTHORS
 Asamiya, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
TITLE
 A large scale structural analysis of cDNAs in a unicellular green
 alga, Chlamydomonas reinhardtii. I. Generation of 3433
 non-redundant expressed sequence tags
JOURNAL
 DNA Res. 6 (6), 369-373 (1999)
MEDLINE
 20152988
PUBMED
 10691129
COMMENT
 Contact: Yasukazu Nakamura
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1332-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
 source
 1..551
 /organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="C9"
 /db_xref="taxon:3055"
 /clone="CM083e05 t"
 /dev_stage="photoautotrophic growth"
 /clone_lib="Chlamydomonas reinhardtii C9"
 /notes="Vector: pBluescriptII SK; Site_1: EcoRI; Site_2:
 XhoI"
 BASE COUNT 94 a 182 c 189 g 85 t 1 others
 ORIGIN
 Query Match 7.2%; Score 40.6; DB 9; Length 551;
 Best Local Similarity 45.3%; Pred. No. 5.1;
 Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
 Qy 42 CGACCTCGTGGGGTACATTCGGTCTGTGGCGCCCCCTTAGGGGGCGTCCAGGGCCCT 101
 Db 108 CGAGCTCATCTGTGTCATTGTGGCGGCACCTGCCAACATGAAGGACGTGCTCAGGACCT 167
 Qy 102 GGCGCATGGCGTCCGGGTTCGTGGAGGACGGCTGAACATGTCACAAAGGGAATTTGCCCGG 161
 Db 168 GCGCGGCGCGCGCGCGAGTGGGNGGCGGCTACGGCGACAGTCCGTGAGCTTGGGCGC 227
 Qy 162 TTGCTCTTTCTATCTTCCTTTGGCTTTGGCTTCTGTCCTGTCACCGTTCACGTTCCGCTCCG 221
 Db 228 CCGCAAGGTGTTTACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCAGAACCCCG 287
 Qy 222 TTATGAATGGCAACACTGTCCGGGTATGACATGTCAAGAACTGTCTCCAACTCAAG 281
 Db 288 CTTGCGCGTCCGCTCGTGGGCGCACTCGTGGGCGGCGCACCGCGGCTGCTGCTGAT 347
 Qy 282 CATTTGTGTATAGGCAGCGGACATGATCATGCAACCCCGGGTGTGCTGCTGCTGCTGCTG 341
 Db 348 CTTGATGCACCAACACGACGAGGAGTTTGGCGCGGATCTACGGCGGCTGTCATGCGCGG 407
 Qy 342 GGAGAACAACTTTCCCGCTCTGGT 368
 Db 408 CAAGAAGAGCAAGGCGCAGCTACATGAT 434
RESULT 7
BI996341
LOCUS
 1031037A07.v2 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION
 Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION
 BI996341
VERSION
 BI996341.1 GI:16431115
KEYWORDS
 EST.
SOURCE
 Chlamydomonas reinhardtii
 Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 552)
Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers
1..552
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda Zap II"
/note="vector; Bluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffery McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. Bluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
93 184 189 86 t

Query Match	7.2%;	Score 40.6;	DB 12;	Length 552;
Best Local Similarity	45.3%;	Pred. No. 5.1;		
Matches 148;	Conservative	0;	Mismatches 179;	Indels 0; Gaps 0;
QY	42	CGACCTGTGGGGTACATTCCGCTCGTGGCGCCCCCTAGGGGCGCTGCCAGGGCCCT	101	
DB	110	CGAGCTCATCTCTGGTCATTGTGCGGGGCACCTCCCAACATGAAGNACGTGCTGACGGACCT	169	
QY	102	GGCGCATGGCGTCCGGGTTCTGGAGGACGGCGTGAACTATGCAACAGAGGAATTTGCCCGG	161	
DB	170	GGCCGGCGCGCGCGAGTGGAGGGCGGCTACGCGCACGAGTCCGTGAGCTTGGGCGC	229	
QY	162	TTGTCTCTTCTCTATCTTCTCTTTGGCTTTGTGTCTGTCTGTACCGTTCCAGCTTCCGC	221	
DB	230	CCGCAAGTGTTTTCACGAGATCAAGGAGTACGTGTGTGAACCTCAAGGCCAGAACCCGAC	289	
QY	222	TTATGAAGTGGCNAACGTGTCCGGATGTACCATGTACAGAACGACTGCTCCAACCTCAAG	281	
DB	290	CTTCGCGCTCGCTGCGTGGGCCACTCGCTGGGGCGGGCAACGCGCGCTGCTGTGAT	349	
QY	282	CATTGTGTATGAGGACGCGGACATGATCATGCAACCCCGGGTGGGTGCCCTTGGCTTCG	341	
DB	350	CTGATGCACCAACGACGAGGATTTGGGCGCGCATCTACGGGCGGCTGCCATGCCGGG	409	
QY	342	GGAGAACAACTCTTCCGCTGCTGGGT	368	
DB	410	CAAGAAGACGAAGGCGACATCATGAT	436	

RESULT 8				
BI727879				
LOCUS	BI727879	584 bp	mRNA	linear
				EST 19-SEP-2001


```

DEFINITION      1031095C12.y1 C. reinhardtii CC-1690, Stress II (normalized),
                Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      B1727879
VERSION        B1727879.1 GI:15703574
KEYWORDS       EST.
SOURCE         Chlamydomonas reinhardtii
ORGANISM       Chlamydomonas reinhardtii
                Chlamydomonas reinhardtii
                Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
                Chlamydomonadaceae; Chlamydomonas.
REFERENCE      1 (bases 1 to 584).
                Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre,
                P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
                Analyses of the Chlamydomonas reinhardtii Genome: A Model,
                Unicellular System for Analyzing Gene Function and Regulation in
                Vascular Plants. Project: 1031
                Unpublished
COMMENT        Contact: Charles Hauser
                DCMB Box 91000
                Duke University
                Durham, NC 27708-1000
                Tel: 919 613 8159
                Fax: 919 613 8177
                Email: chauser@duke.edu.
FEATURES       Location/Qualifiers
                1..584
                /organism="Chlamydomonas reinhardtii"
                /mol_type="mRNA"
                /strain="CC-1690 wild type mt+ 21gr"
                /db_xref="taxon:3055"
                /clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
                ), Lambda Zap II"
                /notes="vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
                XhoI; Stress condition II library, constructed by John
                Davies and Jeffrey McDermott, combines cDNAs from CC-1690
                cells grown to mid-log phase in TAP (NH4+ - containing)
                and shifted to TAP - NO3- (24hrs); #2 production
                conditions (0, 12hr, 24hr) see Meilis et al., (2000) Plant
                Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
                sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
                polyA mRNA was purified from each sample, pooled and cDNA
                synthesized. The cDNA was directionally cloned into lambda
                Zap II (Stratagene) in the EcoRI (5') and XhoRI (3')
                sites. pBluescript II SK- plasmids were excised from the
                lambda Zap clones by superinfection with EXAssist
                (Stratagene) phage. The library was normalized using
                method 4 described in Bonaldo et al., (1996) Genome
                Research 6: 791-806."
BASE COUNT     106 a      188 c      197 g      93 t
ORIGIN
Query Match    .7.2%; Score 40.6; DB 12; Length 584;
Best Local Similarity 45.3%; Pred. No. 5.2;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

Qy 42 CGACCTCGTGGGTACATTCGGTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCT 101
    |||
Db 47 CGAGCTCATCTGTGTCATTGTGCGCGCACTGCCAACATCAAGACGCTGCTGACGACCT 106
    |||

Qy 102 GCGCATGGCTCGGGTTCTGGAGGACGGGTGAATATGCAACAGGAATTGCGCCGG 161
    |||
Db 107 GCGCGGCGCGCGCGGAGTGGGAGGGCGCTAGCGCACAGTCCGTGAGCTTGGGCGC 166
    |||

Qy 162 TTGCTCTTTCTATCTTCTCTTTGCTTTGCTTGCTGCTGCTGACGGTTCCAGCTTCCGC 221
    |||
Db 167 CCGCAAGTGTTCACAGAGATCAAGGAGTAGTGCTGAACCTCAAGGCCAGAACCCAG 226
    |||

Qy 222 TTATGAAGTCGCAACGTGTTCGGGATGTATACCATGTCAAGAACGACTGTCTCAACTCAAG 281
    |||
Db 227 CTTCCGCGTTCGCTGCGTGGCCACATCGCTGGGGCGGGCACCGCGGCTGCCTGCGAT 286
    |||

Qy 282 CATTGTGTATGAGCAGCGGACATGATCATGTCACACCCCGGGTGGCGTCCCTGCGGTTCG 341
    |||
Db 287 CCTGATGCACACGACGAGGAGTTTGGGGCGCGCATCTACGGGCGGTGCCATGCGCGG 346
    |||

```


[illegible]


```

QY 332 CTTGGCTTGGGAGAACAACTCTTCCTGGCTGGGTAGCGGTCAACCCCGACGCTGCGAG 391
Db 832 SGCSKGTTKTKTTTGTGCTGTTGAAGASABAGTWTGAGGGGGGGCCCYCSCMCCCCCCYB 773
QY 392 CTAGGAACAGCCAGCGTCCCCACACGACA 420
Db 772 BBCCCMCHCTKCSCKMCCRGACTYCCCA 744

RESULT 12
BJ252669/c
LOCUS
DEFINITION
BJ252669 Y. Ogiwara unpublished cDNA library, Wh_f Triticum
aestivum cDNA clone whf25g19 3', mRNA sequence.
ACCESSION
BJ252669
VERSION
BJ252669.1 GI:20061830
KEYWORDS
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 359)
REFERENCE
Ogiwara,Y. and Murai,K.
EXpressed genes in Triticum aestivum
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/tissue_type="spike at flowering date"
/dev_stage="Peekes' scale 10.5.1"
/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"
BASE COUNT 70 a 115 c 107 g 67 t
ORIGIN
Query Match 7.0%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CTTACATCGGCTTCGCGACCTCTGGGGTACATTCCGCTCGCGCGCCCGCTAGG 83
Db 297 CTTCAAGTGAACAGCGCGCTCTGGAAGGCGCTCAGGGCGTGCAGCGCGTCCGCTCGG 238
QY 84 GGGCGCTGCAGGCGCTCGCGATCGCGTCCGGTTCTGGAGGAGCGCGTGAACATATG 142
Db 237 GGACGCGCGCAGCGCCCTGGCGCAGGACGTGCGAGTCTGCCGTGCACGTGCCCAAGG 179

RESULT 13
BJ246716
LOCUS
DEFINITION
BJ246716 Y. Ogiwara unpublished cDNA library, Wh_f Triticum
aestivum cDNA clone whf25g19 5', mRNA sequence.
ACCESSION
BJ246716
VERSION
BJ246716.1 GI:20058228
KEYWORDS
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 375)
REFERENCE

```

AUTHORS TITLE JOURNAL COMMENT

Ogiwara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES source

1..375
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Peekes' scale 10.5.1"
/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"
81 a 107 c 110 g 77 t

BASE COUNT 81 a 107 c 110 g 77 t
ORIGIN

Query Match 7.0%; Score 39; DB 12; Length 375;
Best Local Similarity 58.0%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 24 CTTACATCGGCTTCGCGACCTCTGGGGTACATTCCGCTCGCGCGCCCGCTAGG 83
Db 36 CTTCAAGTGAACAGCGCGCTCTGGAAGGCGCTCAGGGCGTGCAGCGCGTCCGCTCGG 95
QY 84 GGGCGCTGCAGGCGCTCGCGATCGCGTCCGGTTCTGGAGGAGCGCGTGAACATATG 142
Db 96 GGACGCGCGCAGCGCCCTGGCGCAGGACGTGCGAGTCTGCCGTGCACGTGCCCAAGG 154

RESULT 14 CC335916 LOCUS

OGUJ60TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0393124,
genomic survey sequence.
CC335916
CC335916.1 GI:30805329
GSS.
Ze mays
Ze mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)

REFERENCE AUTHORS

Whiteley,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

TITLE JOURNAL COMMENT

FEATURES source

1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0393124"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBSCS-; Site:1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

us-09-899-303a-23.rst

Mon Dec 22 13:28:41 2003

Search completed: December 20, 2003, 06:54:49
Job time : 1444.21 secs

```

BASE COUNT      103 a      298 c      291 g      148 t
ORIGIN

Query Match      7.0%; Score 39; DB 29; Length 840;
Best Local Similarity 53.6%; Pred. No. 15;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 39 CGCCGACCTCGTGGGGTACATTCCGCTCGTCGGCGCCCTAGGGGGGCTGCCAGGGC 98
    |||||
Db 292 CCCGGGTCGACGGGCCACTCCGCTTCACTGCGCGGCTCCGGCGGCTCGCGGAAGC 351
    |||||

QY 99 CTTGGCGCATGGGTCGGGTTCTGAGGACGCGTGAACTATGCAACAGGGAATTGGC 158
    |||||
Db 352 TGTGGCGCTGCTCAAGCTGCTCTGCTGGTGGCGGCGCGGTAACCTCCGACTCCGC 411
    |||||

QY 159 CGGTTGCTCTTTCTATCTTCCTCTTCTTGGCT 189
    |||||
Db 412 CGGTCGCTTGCTGCGGACGTCATCTTGTCT 442
    |||||

RESULT 15
CD446071      873 bp      mRNA      linear      EST 03-JUN-2003
DEFINITION    EL01T0207B11.b EndospERM_4 Zea mays cDNA, mRNA sequence.
ACCESSION    CD446071
VERSION      CD446071.1 GI:31361714
KEYWORDS
SOURCE
ORGANISM      Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS      Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
                Messing, J.
TITLE        Sequencing of the maize endospERM ESTs
JOURNAL
COMMENT      Contact: Lai, Jinsheng
                Dr. Joachim Messing's lab
                Waksman Institute, Rutgers University
                190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
                Tel: 732-445-3801
                Fax: 732-445-5735
                Email: jlai@waksman.rutgers.edu
                Seq primer: T3.
FEATURES
source
    1..873
        /organism="Zea mays"
        /mol_type="mRNA"
        /culturivar="W22"
        /db_xref="taxon:4577"
        /tissue_type="EndospERM of 7-23DAP"
        /clone_lib="EndospERM 4"
        /notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                XhoI"

BASE COUNT      109 a      300 c      294 g      170 t
ORIGIN

Query Match      7.0%; Score 39; DB 14; Length 873;
Best Local Similarity 53.6%; Pred. No. 15;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 39 CGCCGACCTCGTGGGGTACATTCCGCTCGTCGGCGCCCTAGGGGGGCTGCCAGGGC 98
    |||||
Db 583 CCCGGGTCGACGGGCCACTCCGCTTCACTGCGCGGCTCCGGCGGCTCGCGGAAGC 642
    |||||

QY 99 CTTGGCGCATGGGTCGGGTTCTGAGGACGCGTGAACTATGCAACAGGGAATTGGC 158
    |||||
Db 643 TGTGGCGCTGCTCAAGCTGCTCTGCTGGTGGCGGCGCGGTAACCTCCGACTCCGC 702
    |||||

QY 159 CGGTTGCTCTTTCTATCTTCCTCTTGGCT 189
    |||||
Db 703 CGGTCGCTTGCTGCGGACGTCATCTTGTCT 733
    |||||

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 40.4591 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-23

Perfect score: 561

Sequence: 1 ATGTTGGTAAAGTCAATCGA.....TGATGATGAACGTGTAATAG 561

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	561	100.0	561	3	US-08-612-973-23
2	561	100.0	561	3	US-08-927-597-23
3	556	99.1	606	3	US-08-612-973-25
4	556	99.1	606	3	US-08-927-597-25
5	556	99.1	636	3	US-08-612-973-27
6	556	99.1	636	3	US-08-927-597-27
7	556	99.1	723	3	US-08-612-973-21
8	556	99.1	723	3	US-08-927-597-21
9	474	84.5	795	3	US-08-612-973-5
10	474	84.5	795	3	US-08-927-597-5
11	471	84.0	2082	3	US-08-612-973-47
12	471	84.0	2082	3	US-08-927-597-47
13	471	84.0	2433	3	US-08-612-973-49
14	471	84.0	2433	3	US-08-927-597-49
15	461.4	82.2	633	3	US-08-612-973-7
16	461.4	82.2	633	3	US-08-927-597-7
17	454	80.9	636	3	US-08-612-973-13
18	454	80.9	636	3	US-08-927-597-13
19	415.4	74.0	1539	2	US-08-470-426B-17
20	415.4	74.0	1863	2	US-08-470-426B-14
21	413.4	73.7	1037	1	US-08-462-195-1
22	413.4	73.7	1037	1	US-08-636-883-1
23	413.4	73.7	1037	3	US-09-127-829-1
24	410.6	73.2	932	1	US-08-081-072-15
25	410.6	73.2	932	1	US-08-449-093A-15
26	410.6	73.2	9595	3	US-09-014-416-4
27	410.6	73.2	9599	3	US-09-014-416-6

28 409 72.9 2116 3 US-08-191-160-21 Sequence 21, Appl
29 403.8 72.0 9472 4 US-08-150-204E-96 Sequence 96, Appl
30 397.4 70.8 1167 1 US-08-324-977-9 Sequence 9, Appl
31 397.4 70.8 1167 2 US-08-384-616-9 Sequence 9, Appl
32 397.4 70.8 1167 2 US-08-904-686A-9 Sequence 9, Appl
33 397.4 70.8 1167 3 US-09-315-850-9 Sequence 9, Appl
34 397.4 70.8 1499 1 US-08-324-977-3 Sequence 3, Appl
35 397.4 70.8 1499 2 US-08-384-616-3 Sequence 3, Appl
36 397.4 70.8 1499 2 US-08-904-686A-3 Sequence 3, Appl
37 397.4 70.8 1499 3 US-09-315-850-3 Sequence 3, Appl
38 397.4 70.8 6039 1 US-08-324-977-11 Sequence 11, Appl
39 397.4 70.8 6039 2 US-08-384-616-11 Sequence 11, Appl
40 397.4 70.8 6039 2 US-08-904-686A-11 Sequence 11, Appl
41 397.4 70.8 6039 3 US-09-315-850-11 Sequence 11, Appl
42 397.4 70.8 9030 1 US-08-324-977-13 Sequence 13, Appl
43 397.4 70.8 9030 2 US-08-384-616-13 Sequence 13, Appl
44 397.4 70.8 9030 2 US-08-904-686A-13 Sequence 13, Appl
45 397.4 70.8 9030 3 US-09-315-850-13 Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-23
; Sequence 23, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..555


```

; NAME/KEY: CDS
; LOCATION: 1..603
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..600
; US-08-927-597-25

Query Match
Best Local Similarity 99.1%; Score 556; DB 3; Length 606;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGGCGCATGGGTCGGGTT 120
DB 61 CCGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGGCGCATGGGTCGGGTT 120
QY 121 CTGGAGACGGCTGAATATGCAACAGGGAATTTGCCGGTTGCTTCTTCTATCTTC 180
DB 121 CTGGAGACGGCTGAATATGCAACAGGGAATTTGCCGGTTGCTTCTTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGCTCTGTGACCGTTCCAGCTTCGCTTATGAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTGCTGCTCTGTGACCGTTCCAGCTTCGCTTATGAGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACG 300
DB 241 TCCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACG 300
QY 301 GACATGATCATGACACACCCCGGTTGGTCCCTGCTTCGGGAGAACAACTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGTTGGTCCCTGCTTCGGGAGAACAACTTCCCGC 360
QY 361 TGCTGGTGGTACGCTACACCCCGGTTGGTCCCTGCTTCGGGAGAACAACTTCCCGC 420
DB 361 TGCTGGTGGTACGCTACACCCCGGTTGGTCCCTGCTTCGGGAGAACAACTTCCCGC 420
QY 421 ATAGACCGCCAGCTGATTTCCAGCTGTTTCCACATCTCGCTTCGCCGCGCATGAGCG 480
DB 421 ATAGACCGCCAGCTGATTTCCAGCTGTTTCCACATCTCGCTTCGCCGCGCATGAGCG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTTGG 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTTGG 540
QY 541 ATGATGATGAAGTGGT 556
DB 541 ATGATGATGAAGTGGT 556

RESULT 5
US-08-612-973-27
; Sequence 27, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUUSE, MARIE-ANGE
; APPLICANT:
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630
; US-08-612-973-27

Query Match
Best Local Similarity 99.1%; Score 556; DB 3; Length 636;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGGCGCATGGGTCGGGTT 120
DB 61 CCGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGGCGCATGGGTCGGGTT 120
QY 121 CTGGAGACGGGCTGAATATGCAACAGGGAATTTGCCGGTTGCTTCTTCTATCTTC 180
DB 121 CTGGAGACGGGCTGAATATGCAACAGGGAATTTGCCGGTTGCTTCTTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGCTCTGTGACCGTTCCAGCTTCGCTTATGAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTGCTGCTCTGTGACCGTTCCAGCTTCGCTTATGAGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACG 300
DB 241 TCCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACG 300
QY 301 GACATGATCATGACACACCCCGGTTGGTCCCTGCTTCGGGAGAACAACTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGTTGGTCCCTGCTTCGGGAGAACAACTTCCCGC 360
QY 361 TGCTGGTGGTACGCTACACCCCGGTTGGTCCCTGCTTCGGGAGAACAACTTCCCGC 420
DB 361 TGCTGGTGGTACGCTACACCCCGGTTGGTCCCTGCTTCGGGAGAACAACTTCCCGC 420
QY 421 ATAGACCGCCAGCTGATTTCCAGCTGTTTCCACATCTCGCTTCGCCGCGCATGAGCG 480
DB 421 ATAGACCGCCAGCTGATTTCCAGCTGTTTCCACATCTCGCTTCGCCGCGCATGAGCG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTTGG 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTTGG 540
QY 541 ATGATGATGAAGTGGT 556
DB 541 ATGATGATGAAGTGGT 556
```


Db 541 ATGATGATGAAGTGGT 556

RESULT 6
US-08-927-597-27
; Sequence 27, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630
US-08-927-597-27

Query Match
Best Local Similarity 99.1%; Score 556; DB 3; Length 636;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTGCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGTGCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CGCTCGTCCGGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGCGCGTCCGGGTT 120
Db 61 CGCTCGTCCGGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGCGCGTCCGGGTT 120
QY 121 CTGGAGAGCGGTGAAGTATGATGCAAGGAAATTCGCCGGTTCCTTCTATCTTC 180
Db 121 CTGGAGAGCGGTGAAGTATGATGCAAGGAAATTCGCCGGTTCCTTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAGTG 240
Db 181 CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAGTG 240
QY 241 TCGGGATGATACCATGTACGACGACGACTCTCCAACTCAAGCATTTGTATGAGGAGCG 300
Db 241 TCGGGATGATACCATGTACGACGACGACTCTCCAACTCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGCCTGCGTTCCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCGTTCCGGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGTAGCGCTACCCCGGAGCTTCGAGCTAGGAAGCCAGCGTCCCGACAGACA 420
Db 361 TCGTGGTAGCGCTACCCCGGAGCTTCGAGCTAGGAAGCCAGCGTCCCGACAGACA 420
QY 421 ATACGACGCGCAGCTCGATTCCCGAGCTGTTCAACATCTCGCTCGCGCATGAGCGGTG 480
Db 421 ATACGACGCGCAGCTCGATTCCCGAGCTGTTCAACATCTCGCTCGCGCATGAGCGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAAGTGGT 556
Db 541 ATGATGATGAAGTGGT 556

RESULT 7
US-08-612-973-21
; Sequence 21, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO

Db 481 CAGACTGCAATTGCTCAATCTATCCGGCCACATACGGGTACCGGTATGGGTGGAT 540
Qy 541 ATGATGATGAAGTGGT 556
Db 541 ATGATGATGAAGTGGT 556
RESULT 9
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789
US-08-612-973-5
Query Match 84.5%; Score 474; DB 3; Length 795;
Best Local Similarity 88.5%; Pred. No. 2.8e-118;
Matches 556; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
Qy 1 ATGCTGGTAAAGTCATCGATCCCTACATCGGCTTCGCCACCTCGTGGGGTACATT 60
Db 1 ATGTTGGTAAAGTCATCGATCCCTACATCGGCTTCGCCACCTCGTGGGGTACATT 60
Qy 61 CCCTCGTCCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTT 120
Db 61 CCCTCGTCCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTT 120
Qy 121 CTGAGAGCGCGGTGAACATATGCAACAGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGAGAGCGCGGTGAACATATGCAACAGGAATTTGCCCGGTTGCTCTTCTATCTTC 180
Qy 181 CTCTTGGCTTTGCTGTCTGACCGTTTCAGCTTTCGGCTTATGAAGTCCGCAACGTG 240
Db 181 CTCTTGGCTTTGCTGTCTGACCGTTTCAGCTTTCGGCTTATGAAGTCCGCAACGTG 240
Qy 241 TCCGGGATGTACCATGTACAGAACGATGCTGCCAATCAAGCAATTTGTATGAGGACGG 300
Db 241 TCCGGGATGTACCATGTACAGAACGATGCTGCCAATCAAGCAATTTGTATGAGGACGG 300
Qy 301 GACATGATCATGCACACCCCGGTCGCTCGCTTCCGAGTTCGGGAGAACACTTCTCCGC 360
Db 301 GACATGATCATGCACACCCCGGTCGCTCGCTTCCGAGTTCGGGAGAACACTTCTCCGC 360
Qy 361 TGCTGGGTAGCGCTCACCCCGGTCGCTCGCTTCCGAGTTCGGGAGAACACTTCTCCGC 420
Db 361 TGCTGGGTAGCGCTCACCCCGGTCGCTCGCTTCCGAGTTCGGGAGAACACTTCTCCGC 420
Qy 421 ATACGACGCCACGTCGAT----- 438
Db 421 ATACGACGCCACGTCGATTTGCTTGGGGCGGCTTCTTCTGTTCCGCTATGTACGTG 480
Qy 439 -----TCCAGCTGTTACCATCTCGCTCGCGG 468
Db 481 GGGGACCTCTGGGATCTGTTCTCGTCTCCAGCTGTTTCCAGCTCTCGCTCGCGG 540
Qy 469 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAAGCGGTACCGT 528
Db 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAAGCGGTACCGT 600
Qy 529 ATGGCTTGGGATATGATGATGAAGTGGT 556
Db 601 ATGGCTTGGGATATGATGATGAAGTGGT 628
RESULT 10
US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE: 11-MAR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000

us-09-899-303a-23.rni

Mon Dec 22 13:28:39 2003

```

;
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789
;
; US-08-927-597-5

```

```

Query Match      84.5%; Score 474; DB 3; Length 795;
Best Local Similarity 88.5%; Pred. No. 2.8e-118; Indels 72; Gaps 1;
Matches 556; Conservative 0; Mismatches 0;

QY 1 ATGTTGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCACCTCGTGGGTACATT 60
QY 61 CCGTCGTGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGCGCATGGCGTCCGGGTT 120
DB 61 CCGTCGTGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGCGCATGGCGTCCGGGTT 120
QY 121 CTGGAGGACGCGGTGAATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTC 180
DB 121 CTGGAGGACGCGGTGAATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAAGTG 240
DB 181 CTCTTGGCTTTGCTGCTGTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAAGTG 240
QY 241 TCCGGGATGACCATGTACGAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGACCATGTACGAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATCATCACACCCCGGGTGGTGGCTCGCTCGGTTCGGGAGAACACTTCCCGC 360
DB 301 GACATGATCATCACACCCCGGGTGGTGGCTCGCTCGGTTCGGGAGAACACTTCCCGC 360
QY 361 TGTGGGTAGCGCTACCCCGGCTCGAGCTAGGAAAGCGCAGCGTCCCGACCAACGACA 420
DB 361 TGTGGGTAGCGCTACCCCGGCTCGAGCTAGGAAAGCGCAGCGTCCCGACCAACGACA 420
QY 421 ATACGACGCGACGTCGAT----- 438
DB 421 ATACGACGCGACGTCGATTTGCTCGTTGGGGCGGTGCTTTCTGTTCCGCTATGTACGTG 480
QY 439 -----TCCAGCTGTTCCACCATCTCGCTCGCGG 468
DB 439 -----TCCAGCTGTTCCACCATCTCGCTCGCGG 468
QY 481 GGGGACCTCTCGCGGATCTGTTCCTGCTCTCCAGCTGTTTCCAGCTCTCGCGG 540
DB 481 GGGGACCTCTCGCGGATCTGTTCCTGCTCTCCAGCTGTTTCCAGCTCTCGCGG 540
QY 469 CATGAGACGCTGAGGAGTCAATTTGCTCAATCTATCCCGGCAACATAACGGGTACCGT 528
DB 469 CATGAGACGCTGAGGAGTCAATTTGCTCAATCTATCCCGGCAACATAACGGGTACCGT 528
QY 541 CATGAGACGCTGAGGAGTCAATTTGCTCAATCTATCCCGGCAACATAACGGGTACCGT 600
DB 541 CATGAGACGCTGAGGAGTCAATTTGCTCAATCTATCCCGGCAACATAACGGGTACCGT 600
QY 529 ATGCTTTGGGATATGATGATGAATGGT 556
DB 529 ATGCTTTGGGATATGATGATGAATGGT 556
QY 601 ATGCTTTGGGATATGATGATGAATGGT 628
DB 601 ATGCTTTGGGATATGATGATGAATGGT 628

```

```

RESULT 11
US-08-612-973-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT

```

```

;
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYS, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
;
; US-08-612-973-47

```

```

Query Match      84.0%; Score 471; DB 3; Length 2082;
Best Local Similarity 88.5%; Pred. No. 2.4e-117; Indels 72; Gaps 1;
Matches 553; Conservative 0; Mismatches 0;

QY 4 TTGGGTAAGGTCAATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 63
DB 4 TTGGGTAAGGTCAATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 63
QY 64 CTGTCGGCGCCCGCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGCGTTCGGGTTCTG 123
DB 64 CTGTCGGCGCCCGCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGCGTTCGGGTTCTG 123
QY 124 GAGGACGCGGTGAACATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTCCTC 183
DB 124 GAGGACGCGGTGAACATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTCCTC 183
QY 184 TTGCTTTGCTGCTGCTGACCGTTCCAGCTTCCGCTTTATGAAGTGGCGCAACGTTCC 243
DB 184 TTGCTTTGCTGCTGCTGACCGTTCCAGCTTCCGCTTTATGAAGTGGCGCAACGTTCC 243
QY 244 GGGATGTACCAATGTACGAAACGACTGTCCAACTCAAGCATTTGTGTATGAGGACGCGAC 303
DB 244 GGGATGTACCAATGTACGAAACGACTGTCCAACTCAAGCATTTGTGTATGAGGACGCGAC 303
QY 304 ATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCCCGCTGC 363
DB 304 ATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCCCGCTGC 363

```


QY 364 TGGGTAGCGCTACCCCGCAGCTCGCAGCTAGGAAAGCCGAGCGTCCCGCCACGACAATA 423
Db 364 TGGGTAGCGCTACCCCGCAGCTCGCAGCTAGGAAAGCCGAGCGTCCCGCCACGACAATA 423
QY 424 CGACGCCAGCTCGAT----- 438
Db 424 CGACGCCAGCTCGATTTGCTGCTGGGGGGGCTGCTTCTGTTCCGCTATGATGAGTGGGG 483
QY 439 -----TCCAGCTGTTCACCATCTCGCCTCGCCGGGCAT 471
Db 484 GACCTCTGCGGATCTGTTCTCTCCGCTCCAGCTGTTACCATCTCGCCTCGCCGGCAT 543
QY 472 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG 531
Db 544 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG 603
QY 532 GCTTGGGATATGATGAACTGGT 556
Db 604 GCTTGGGATATGATGAACTGGT 628

RESULT 12

US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:

; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-927-597-47

Query Match 84.0%; Score 471; DB 3; Length 2082;
Best Local Similarity 88.5%; Pred. No. 2.4e-117; Indels 72; Gaps 1;
Matches 553; Conservative 0; Mismatches 0;
QY 4 TTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCCG 63
Db 4 TTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCCG 63
QY 64 CTCGTCCGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTCTTG 123
Db 64 CTCGTCCGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTCTTG 123
QY 124 GAGGACGGGTGAACATATCAACAGGGAAATTTGCCCGGTTGCTTCTCTATCTTCCTC 183
Db 124 GAGGACGGGTGAACATATCAACAGGGAAATTTGCCCGGTTGCTTCTCTATCTTCCTC 183
QY 184 TTGGCTTTGCTGTCTGTCTGACCGCTTCAGAGTTCGCGCTTATGAAGTGGCGCAACGTGTC 243
Db 184 TTGGCTTTGCTGTCTGTCTGACCGCTTCAGAGTTCGCGCTTATGAAGTGGCGCAACGTGTC 243
QY 244 GGGATGTACCATGTTCACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGCGGAC 303
Db 244 GGGATGTACCATGTTCACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGCGGAC 303
QY 304 ATGATCATGCACACCCCGGGTGGCGCTCGGTTCCGGGAGAACAACTCTTCCCGCTGC 363
Db 304 ATGATCATGCACACCCCGGGTGGCGCTCGGTTCCGGGAGAACAACTCTTCCCGCTGC 363
QY 364 TGGGTAGCGCTCACCCCGCAGCTCGCAGCTAGGAAGCCAGCGTCCCGCCACGACAATA 423
Db 364 TGGGTAGCGCTCACCCCGCAGCTCGCAGCTAGGAAGCCAGCGTCCCGCCACGACAATA 423
QY 424 CGACGCCAGCTCGAT----- 438
Db 424 CGACGCCAGCTCGATTTGCTGCTGGGGGGGCTGCTTCTGTTCCGCTATGATGAGTGGGG 483
QY 439 -----TCCAGCTGTTCACCATCTCGCCTCGCCGGGCAT 471
Db 484 GACCTCTGCGGATCTGTTCTCTCCAGCTGTTACCATCTCGCCTCGCCGGCAT 543
QY 472 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG 531
Db 544 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG 603
QY 532 GCTTGGGATATGATGAACTGGT 556
Db 604 GCTTGGGATATGATGAACTGGT 628

RESULT 13

US-08-612-973-49
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICANT: MAERTENS, GEERT
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2430
NAME/KEY: mat_peptide
LOCATION: 1..2427
US-08-612-973-49

Query Match 84.0%; Score 471; DB 3; Length 2433;
Best Local Similarity 88.5%; Pred. No. 2.5e-117;
Matches 553; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 4 TTGGGTAAGGTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCG 63
Db 355 TTGGGTAAGGTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCG 414
QY 64 CTCGTCGGCGCCCTAGGGGGCTGCCAGGCGCTTCGCGATGGCGTCCGGTTCG 123
Db 415 CTCGTCGGCGCCCTAGGGGGCTGCCAGGCGCTTCGCGATGGCGTCCGGTTCG 474
QY 124 GAGGACGGCGTGAATATGCAACAGGGAATTGGCGGCTTCCTTCTCTATCTTCTC 183
Db 475 GAGGACGGCGTGAATATGCAACAGGGAATTGGCGGCTTCCTTCTCTATCTTCTC 534
QY 184 TTGGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
Db 535 TTGGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
QY 244 GGGATGTACCATGTCAACAGCACTGCTTCCAACTCAAGCAATGTGTATGAGGACGGAC 303
Db 595 GGGATGTACCATGTCAACAGCACTGCTTCCAACTCAAGCAATGTGTATGAGGACGGAC 654
QY 304 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 655 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
QY 364 TGGGTAGGCTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
Db 715 TGGGTAGGCTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774
QY 424 CGACGCCACGTGCAT----- 438
Db 775 CGACGCCACGTGCAT----- 834
QY 439 -----TCCAGCTGTTCACCATCTCGGCTCGCGGGCAT 471
Db 835 GACCTCTCGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 894
QY 472 GAGACGGTGCAGGACTGCAATTGCTCAATTATCTCCCGGCCACATAACGGGTACCGTATG 531

Db 895 GAGACGGTGCAGGACTGCAATTGCTCAATTATCTCCCGGCCACATAACGGGTACCGTATG 954
QY 532 GCTTGGGATATGATGATGAACCTGGT 556
Db 955 GCTTGGGATATGATGATGAACCTGGT 979
RESULT 14
US-08-927-597-49
Sequence 49, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2430
NAME/KEY: mat_peptide
LOCATION: 1..2427
US-08-927-597-49

Query Match 84.0%; Score 471; DB 3; Length 2433;
Best Local Similarity 88.5%; Pred. No. 2.5e-117;
Matches 553; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 4 TTGGGTAAGGTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCG 63
Db 355 TTGGGTAAGGTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCG 414
QY 64 CTCGTCGGCGCCCTAGGGGGCTGCCAGGCGCTTCGCGATGGCGTCCGGTTCG 123

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2294.49 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-25
Perfect score: 606
Sequence: 1 ATGTTGGTAAGTCAATCGA.....TGCTCCGATCCTCTAATAG 606

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_man.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	606	100.0	606	6	A48687	A48687 Sequence 25
2	606	100.0	606	6	AR157338	AR157338 Sequence
3	606	100.0	606	6	AX452774	AX452774 Sequence
4	606	100.0	606	6	AX685026	AX685026 Sequence
5	598.2	98.7	723	6	A48683	A48683 Sequence 21
6	598.2	98.7	723	6	AR157336	AR157336 Sequence
7	598.2	98.7	723	6	AX452770	AX452770 Sequence
8	598.2	98.7	723	6	AX685022	AX685022 Sequence
9	597.4	98.6	636	6	A48689	A48689 Sequence 27
10	597.4	98.6	636	6	AR157339	AR157339 Sequence
11	597.4	98.6	636	6	AX452776	AX452776 Sequence
12	597.4	98.6	636	6	AX685028	AX685028 Sequence
13	556	91.7	561	6	A48685	A48685 Sequence 23
14	556	91.7	561	6	AR157337	AR157337 Sequence
15	556	91.7	561	6	AX452772	AX452772 Sequence
16	556	91.7	561	6	AX685024	AX685024 Sequence
17	516.2	85.2	795	6	A48667	A48667 Sequence 5
18	516.2	85.2	795	6	AR157325	AR157325 Sequence
19	516.2	85.2	795	6	AX452754	AX452754 Sequence
20	516.2	85.2	795	6	AX685006	AX685006 Sequence
21	513.2	84.7	2082	6	A48709	A48709 Sequence 47
22	513.2	84.7	2082	6	AR157350	AR157350 Sequence
23	513.2	84.7	2082	6	AX452796	AX452796 Sequence
24	513.2	84.7	2082	6	AX685048	AX685048 Sequence
25	513.2	84.7	2433	6	A48711	A48711 Sequence 49
26	513.2	84.7	2433	6	AR157351	AR157351 Sequence
27	513.2	84.7	2433	6	AX452798	AX452798 Sequence
28	513.2	84.7	2433	6	AX685050	AX685050 Sequence
29	460.4	76.0	3296	14	AB008441	AB008441 Hepatitis
30	460.4	76.0	3296	14	AB008442	AB008442 Hepatitis
31	460.4	76.0	3296	14	AB008443	AB008443 Hepatitis
32	460.4	76.0	9379	14	AF207766	AF207766 Hepatitis
33	457.2	75.4	9379	14	AF165052	AF165052 Hepatitis
34	457.2	75.4	9418	14	HCVI32996	AJ132996 Hepatitis
35	456.4	75.3	633	6	A48669	A48669 Sequence 7
36	456.4	75.3	633	6	AR157326	AR157326 Sequence
37	456.4	75.3	633	6	AX452756	AX452756 Sequence
38	456.4	75.3	633	6	AX685008	AX685008 Sequence
39	455.6	75.2	3296	14	AB008445	AB008445 Hepatitis
40	455.6	75.2	3296	14	AB008446	AB008446 Hepatitis
41	455.6	75.2	9344	14	AB049096	AB049096 Hepatitis
42	455.6	75.2	9379	14	HCVPOLYP	AJ000009 Hepatitis
43	455.6	75.2	9386	14	AF165056	AF165056 Hepatitis
44	454	74.9	1615	14	HPCNS1SRPJ	M74813 Hepatitis C
45	454	74.9	3296	14	AB008447	AB008447 Hepatitis

ALIGNMENTS

RESULT 1
A48687
LOCUS A48687
DEFINITION Sequence 25 from Patent WO9604385.
ACCESSION A48687
VERSION A48687.1 GI:2302400
KEYWORDS
SOURCE
ORGANISM
unidentified
unclassified
unclassified
REFERENCE 1 (bases 1 to 606)
Maertens, G., Bosman, F., De, M.G. and Buyse, M.
PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 25 15-FEB-1996;

Mon Dec 22 13:28:41 2003

LOCUS	AR157338	606 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 25 from patent US 6245503.				
ACCESSION	AR157338				
VERSION	AR157338.1	GI:16218271			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 606)				
AUTHORS	Maertens, G., Bosman, F., De Martynoff, G. and Buysse, M.-A.				
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use				
JOURNAL	Patent: US 6245503-A 25 12-JUN-2001;				
FEATURES	Location/Qualifiers				
source	1..606				
BASE COUNT	109 a	193 c	167 g	137 t	
ORIGIN	/organism="unknown"				
	100.0%; Score 606; DB 6; Length 606;				
Query Match	Best Local Similarity 100.0%; Pred. No. 1.3e-123;				
Matches	606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60				
Db	1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60				
Qy	61 CCGCTCTGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120				
Db	61 CCGCTCTGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120				
Qy	121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTGTCTTCTCTATCTTC 180				
Db	121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTGTCTTCTCTATCTTC 180				
Qy	181 CTCCTTGGCTTTGCTCTCTGCTGACCGTTCCAGCTTCGAGCTTCGAGTGCAGCAACGTG 240				
Db	181 CTCCTTGGCTTTGCTCTCTGCTGACCGTTCCAGCTTCGAGCTTCGAGTGCAGCAACGTG 240				
Qy	241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300				
Db	241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300				
Qy	301 GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC 360				
Db	301 GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC 360				
Qy	361 TGCTGGGTAGCGCTCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC 420				
Db	361 TGCTGGGTAGCGCTCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC 420				
Qy	421 ATACGACGGCAGCGTCCGATTTCCAGCTGTTTCCAGCTTCGCTTCGGGAGAACAACTCTTCCCGC 480				
Db	421 ATACGACGGCAGCGTCCGATTTCCAGCTGTTTCCAGCTTCGCTTCGGGAGAACAACTCTTCCCGC 480				
Qy	481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTCCACCGTATGGCTTGGGAT 540				
Db	481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTCCACCGTATGGCTTGGGAT 540				
Qy	541 ATGATGATGAATGCTCAATCTATCCCGGCCACATAACCGGTCCACCGTATGGCTTGGGAT 600				
Db	541 ATGATGATGAATGCTCAATCTATCCCGGCCACATAACCGGTCCACCGTATGGCTTGGGAT 600				
Qy	601 TAATAG 606				
Db	601 TAATAG 606				
RESULT 3					
AX452774					
LOCUS					
Sequence 25 from Patent EP1211315.					
DEFINITION					
ACCESSION					
AX452774					
PAT 06-JUL-2002					

LOCUS	AR157338	606 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 25 from patent US 6245503.				
ACCESSION	AR157338				
VERSION	AR157338.1	GI:16218271			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 606)				
AUTHORS	Maertens, G., Bosman, F., De Martynoff, G. and Buysse, M.-A.				
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use				
JOURNAL	Patent: US 6245503-A 25 12-JUN-2001;				
FEATURES	Location/Qualifiers				
source	1..606				
BASE COUNT	109 a	193 c	167 g	137 t	
ORIGIN	/organism="unknown"				
	100.0%; Score 606; DB 6; Length 606;				
Query Match	Best Local Similarity 100.0%; Pred. No. 1.3e-123;				
Matches	606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60				
Db	1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60				
Qy	61 CCGCTCTGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120				
Db	61 CCGCTCTGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120				
Qy	121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTGTCTTCTCTATCTTC 180				
Db	121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTGTCTTCTCTATCTTC 180				
Qy	181 CTCCTTGGCTTTGCTCTCTGCTGACCGTTCCAGCTTCGAGCTTCGAGTGCAGCAACGTG 240				
Db	181 CTCCTTGGCTTTGCTCTCTGCTGACCGTTCCAGCTTCGAGCTTCGAGTGCAGCAACGTG 240				
Qy	241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300				
Db	241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300				
Qy	301 GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC 360				
Db	301 GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC 360				
Qy	361 TGCTGGGTAGCGCTCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC 420				
Db	361 TGCTGGGTAGCGCTCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC 420				
Qy	421 ATACGACGGCAGCGTCCGATTTCCAGCTGTTTCCAGCTTCGCTTCGGGAGAACAACTCTTCCCGC 480				
Db	421 ATACGACGGCAGCGTCCGATTTCCAGCTGTTTCCAGCTTCGCTTCGGGAGAACAACTCTTCCCGC 480				
Qy	481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTCCACCGTATGGCTTGGGAT 540				
Db	481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTCCACCGTATGGCTTGGGAT 540				
Qy	541 ATGATGATGAATGCTCAATCTATCCCGGCCACATAACCGGTCCACCGTATGGCTTGGGAT 600				
Db	541 ATGATGATGAATGCTCAATCTATCCCGGCCACATAACCGGTCCACCGTATGGCTTGGGAT 600				
Qy	601 TAATAG 606				
Db	601 TAATAG 606				
RESULT 3					
AX452774					
LOCUS					
Sequence 25 from Patent EP1211315.					
DEFINITION					
ACCESSION					
AX452774					
PAT 06-JUL-2002					


```
VERSION      AX452774.1  GI:21712459
KEYWORDS
SOURCE       Hepatitis C virus
ORGANISM     Hepatitis C virus
              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.
TITLE        Recombinant vectors for producing hcv envelope proteins
JOURNAL      Patent: EP 1211315-A 25 05-JUN-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
              source
                1..606
                /organism="Hepatitis C virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:11103"
              CDS
                1..603
                /notes="unnamed protein product"
                /codon_start=1
                /protein_id="CAD38074.1"
                /translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGAAALAHGVRLDGG
                VNYATGNLPGCSFIFLLALLSCLTVPASAYEVNYSVMVHTNDCNSISIVVEADM
                IMHTPCVPCVRENSSRCWVALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
                QDCNCSLYPGHITGHRMNDMMNWSPTTALVVSQRLRL"
              mat_peptide
                1..600
                /product="unnamed"
BASE COUNT   109 a 193 c 167 g 137 t
ORIGIN
Query Match 100.0%; Score 606; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.3e-123;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAAGTCATCGATACCTTACATGGCGCTTCGCCACCTCGGGGATACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCTTACATGGCGCTTCGCCACCTCGGGGATACATT 60
QY 61 CCCTCGTCCGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTT 120
DB 61 CCCTCGTCCGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAAGGAAATTTGCCGGTTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAAGGAAATTTGCCGGTTGCTCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGACCGTTCCAGCTTCGCGTTCGAGTTCGAGTTCGCGCAACG 240
DB 181 CTCTTGGCTTTGCTGCTCTGACCGTTCCAGCTTCGCGTTCGAGTTCGCGCAACG 240
QY 241 TCCGGATGTPACCATGTCAAGACGACTGCTCCAACTCAAGCAATTTGTATGAGGACGCG 300
DB 241 TCCGGATGTPACCATGTCAAGACGACTGCTCCAACTCAAGCAATTTGTATGAGGACGCG 300
QY 301 GACATGATCATGACACCCCGGGTGGTGGTTCGCGTTCGAGTTCGAGTTCGAGTTCGAGTTC 360
DB 301 GACATGATCATGACACCCCGGGTGGTGGTTCGCGTTCGAGTTCGAGTTCGAGTTCGAGTTC 360
QY 361 TGCTGGGTAGCGGTCAACCGCGCTTCGCGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 420
DB 361 TGCTGGGTAGCGGTCAACCGCGCTTCGCGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 420
RESULT 4
AX452774.1  GI:21712459
LOCUS       AX685026
DEFINITION  Sequence 25 from Patent WO2055548.
ACCESSION   AX685026
VERSION     AX685026.1  GI:29371431
KEYWORDS
SOURCE      Hepatitis C virus
ORGANISM    Hepatitis C virus
              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens, G., Bosman, F. and Buyse, M.A.
TITLE        Purified Hepatitis C Virus envelope proteins for diagnostic and
              therapeutic use
JOURNAL      Patent: WO 02055548-A 25 18-JUL-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
              source
                1..606
                /organism="Hepatitis C virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:11103"
              CDS
                1..603
                /notes="unnamed protein product"
                /codon_start=1
                /protein_id="CAD86527.1"
                /db_xref="GI:29371432"
                /translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGAAALAHGVRLDGG
                VNYATGNLPGCSFIFLLALLSCLTVPASAYEVNYSVMVHTNDCNSISIVVEADM
                IMHTPCVPCVRENSSRCWVALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
                QDCNCSLYPGHITGHRMNDMMNWSPTTALVVSQRLRL"
              mat_peptide
                1..600
                /product="unnamed"
BASE COUNT   109 a 193 c 167 g 137 t
ORIGIN
Query Match 100.0%; Score 606; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.3e-123;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAAGTCATCGATACCTTACATGGCGCTTCGCCACCTCGGGGATACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCTTACATGGCGCTTCGCCACCTCGGGGATACATT 60
QY 61 CCCTCGTCCGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTT 120
DB 61 CCCTCGTCCGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAAGGAAATTTGCCGGTTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAAGGAAATTTGCCGGTTGCTCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGACCGTTCCAGCTTCGCGTTCGAGTTCGAGTTCGCGCAACG 240
DB 181 CTCTTGGCTTTGCTGCTCTGACCGTTCCAGCTTCGCGTTCGAGTTCGCGCAACG 240
QY 241 TCCGGATGTPACCATGTCAAGACGACTGCTCCAACTCAAGCAATTTGTATGAGGACGCG 300
DB 241 TCCGGATGTPACCATGTCAAGACGACTGCTCCAACTCAAGCAATTTGTATGAGGACGCG 300
QY 301 GACATGATCATGACACCCCGGGTGGTGGTTCGCGTTCGAGTTCGAGTTCGAGTTCGAGTTC 360
DB 301 GACATGATCATGACACCCCGGGTGGTGGTTCGCGTTCGAGTTCGAGTTCGAGTTCGAGTTC 360
QY 361 TGCTGGGTAGCGGTCAACCGCGCTTCGCGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 420
DB 361 TGCTGGGTAGCGGTCAACCGCGCTTCGCGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 420
```


QY	421	ATACGAGCCAGCTCGATTCCAGGTGTTCCACATCTCGCTCGCGCATGAGACGGTG	480
Db	421	ATACGAGCCAGCTCGATTCCAGGTGTTCCACATCTCGCTCGCGCATGAGACGGTG	480
QY	481	CAGGACTCAATGCTCAATCTATCCCGCCACATAAAGGGTCAACGGTATGGCTTGGGAT	540
Db	481	CAGGACTCAATGCTCAATCTATCCCGCCACATAAAGGGTCAACGGTATGGCTTGGGAT	540
QY	541	ATGATGATGAATGCTCGCTACAAACGGCCCTGGTGGTATCCGACGCTGCTCGGATCCTC	600
Db	541	ATGATGATGAATGCTCGCTACAAACGGCCCTGGTGGTATCCGACGCTGCTCGGATCCTC	600
QY	601	TAATAG 606	
Db	601	TAATAG 606	
RESULT 5			
LOCUS	A48683	723 bp	DNA linear PAT 07-MAR-1997
DEFINITION	Sequence 21 from Patent WO9604385.		
ACCESSION	A48683		
VERSION	A48683.1	GI:2302396	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 723)		
AUTHORS	Maertens, G., Bosman, F., De, M. G. and Buyse, M.		
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE		
JOURNAL	INNOGENETICS NV (BE)		
COMMENT	Patent: WO 9604385-A 21 15-FEB-1996; Other publication CA 2172273 960215 Other publication AU 3382495 960304.		
FEATURES	Location/Qualifiers		
source	1..723		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
CDS	1..720		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAA03134.1"		
	/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGHAARALAHGVRVLEDG VNATGNLPFGCSFSIFLALLSCLTPASAVEVRNVMGYHVTNDCNSNIVYEAADM INHTGCVPCVRNENSSRCWALPTLAARNASVPTTIRRHVDSQLFTSPRRHETV QDNCISIVPGHITGHRMAWDMNWSPTTALVSQLLRIPQAVDVMVAGAHGVLGL AYYSVMGNWAKVLIIVMLLPAP"		
mat_peptide	1..717		
	/product="unnamed"		
BASE COUNT	126 a 220 c 208 g 169 t		
ORIGIN			
Query Match	98.7%;	Score 598.2;	DB 6; Length 723;
Best Local Similarity	99.5%;	Pred. No. 6.8e-122;	
Matches	600;	Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	1	ATGTTGGGTAAAGTCAATCGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAAGTCAATCGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT	60
QY	61	CCGCTCGTGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGCGTTCGGGTT	120
Db	61	CCGCTCGTGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGCGTTCGGGTT	120
QY	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCGCGGTTCCTTCTCTATCTTC	180
Db	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCGCGGTTCCTTCTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGCTCTGTCGACCGGTTCCAGCTTCCCGCTTATGAAGTGGCAACGTG	240
Db	181	CTCTTGGCTTTGCTGCTCTGTCGACCGGTTCCAGCTTCCCGCTTATGAAGTGGCAACGTG	240

QY	421	ATACGAGCCAGCTCGATTCCAGGTGTTCCACATCTCGCTCGCGCATGAGACGGTG	480
Db	421	ATACGAGCCAGCTCGATTCCAGGTGTTCCACATCTCGCTCGCGCATGAGACGGTG	480
QY	481	CAGGACTCAATGCTCAATCTATCCCGCCACATAAAGGGTCAACGGTATGGCTTGGGAT	540
Db	481	CAGGACTCAATGCTCAATCTATCCCGCCACATAAAGGGTCAACGGTATGGCTTGGGAT	540
QY	541	ATGATGATGAATGCTCGCTACAAACGGCCCTGGTGGTATCCGACGCTGCTCGGATCCTC	600
Db	541	ATGATGATGAATGCTCGCTACAAACGGCCCTGGTGGTATCCGACGCTGCTCGGATCCTC	600
QY	601	TAATAG 606	
Db	601	TAATAG 606	
RESULT 5			
LOCUS	A48683	723 bp	DNA linear PAT 07-MAR-1997
DEFINITION	Sequence 21 from Patent WO9604385.		
ACCESSION	A48683		
VERSION	A48683.1	GI:2302396	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 723)		
AUTHORS	Maertens, G., Bosman, F., De, M. G. and Buyse, M.		
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE		
JOURNAL	INNOGENETICS NV (BE)		
COMMENT	Patent: WO 9604385-A 21 15-FEB-1996; Other publication CA 2172273 960215 Other publication AU 3382495 960304.		
FEATURES	Location/Qualifiers		
source	1..723		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
CDS	1..720		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAA03134.1"		
	/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGHAARALAHGVRVLEDG VNATGNLPFGCSFSIFLALLSCLTPASAVEVRNVMGYHVTNDCNSNIVYEAADM INHTGCVPCVRNENSSRCWALPTLAARNASVPTTIRRHVDSQLFTSPRRHETV QDNCISIVPGHITGHRMAWDMNWSPTTALVSQLLRIPQAVDVMVAGAHGVLGL AYYSVMGNWAKVLIIVMLLPAP"		
mat_peptide	1..717		
	/product="unnamed"		
BASE COUNT	126 a 220 c 208 g 169 t		
ORIGIN			
Query Match	98.7%;	Score 598.2;	DB 6; Length 723;
Best Local Similarity	99.5%;	Pred. No. 6.8e-122;	
Matches	600;	Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	1	ATGTTGGGTAAAGTCAATCGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAAGTCAATCGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT	60
QY	61	CCGCTCGTGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGCGTTCGGGTT	120
Db	61	CCGCTCGTGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGCGTTCGGGTT	120
QY	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCGCGGTTCCTTCTCTATCTTC	180
Db	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCGCGGTTCCTTCTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGCTCTGTCGACCGGTTCCAGCTTCCCGCTTATGAAGTGGCAACGTG	240


```
QY 241 TCCGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGC 300
Db 241 TCCGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGC 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTCGCTTCCGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTCGCTTCCGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGCTACCCCGACGCTCGCAGCTAGGAAACCCAGCGTCCCGACGACA 420
Db 361 TGTGGGTAGCGCTACCCCGACGCTCGCAGCTAGGAAACCCAGCGTCCCGACGACA 420
QY 421 ATACGAGCGACGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGTG 480
Db 421 ATACGAGCGACGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTCAACGGTATGGCTTGGAT 540
Db 481 CAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTCAACGGTATGGCTTGGAT 540
QY 541 ATGATGATGAACCTGGTGGCTCAACAGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCTC 600
Db 541 ATGATGATGAACCTGGTGGCTCAACAGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCTC 600
QY 601 TAA 603
Db 601 CAA 603

RESULT 7
AX452770
LOCUS AX452770 723 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 21 from Patent EP1211315.
ACCESSION AX452770
VERSION AX452770.1 GI:21712455
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
          Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, P., de Martynoff, G. and Buyse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 21 05-JUN-2002;
INNOCENTICS N.V. (BE)
FEATURES
    source
        1..723
            /organism="Hepatitis C virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:11103"
        1..720
            /notes="unnamed protein product"
            /protein_id="CAD38072.1"
            /db_xref="GI:21712456"
            /translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLDGG
            VNYATGNLPCGSPSIFELLALLSCLTVPASAYEVNRYVMHVTNDCSNSIYVEAADM
            IMHTPCVPCVRENSSRCWVALTPTLAARNASVPTTIRRHVDSOLFISPRRHETV
            QDCNCSIYPGHITGHMAMDMNWSPTTALVVSQLLRIPQAVVDMVAGHVGVLGAGL
            AYSVMGVNKAUULVLMFLAP"
    mat_peptide
        1..717
            /product="unnamed"
            /base_count 126 a 220 c 208 g 169 t
    CDS
        ORIGIN
            Query Match 98.7%; Score 598.2; DB 6; Length 723;
            Best Local Similarity 99.5%; Pred. No. 6.8e-122;
            Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAAGTCATCGTACCCCTACATCGGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCATCGTACCCCTACATCGGGCTTCGCCGACCTCGTGGGTACATT 60
```

```
QY 61 CCCTCGTGGCGCCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGCTCCGGTT 120
Db 61 CCCTCGTGGCGCCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGCGCTCCGGTT 120
QY 121 CTGAGAGCGGGCGTGAACATATGCAACAGAGAAATTTCCCGGTTGCTCTTCTATCTTC 180
Db 121 CTGAGAGCGGGCGTGAACATATGCAACAGAGAAATTTCCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGCTGTGACCGTTTCCAGCTTTCGCTTATGAAGTGCACAGTG 240
Db 181 CTCTTGGCTTTGCTGCTCTGCTGTGACCGTTTCCAGCTTTCGCTTATGAAGTGCACAGTG 240
QY 241 TCCGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGC 300
Db 241 TCCGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGC 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTCGCTTCCGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTCGCTTCCGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGCTCAACCCCGACGCTCGCAGCTAGGAAACCCAGCGTCCCGACGACA 420
Db 361 TGTGGGTAGCGCTCAACCCCGACGCTCGCAGCTAGGAAACCCAGCGTCCCGACGACA 420
QY 421 ATACGAGCGACGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGTG 480
Db 421 ATACGAGCGACGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTCAACGGTATGGCTTGGAT 540
Db 481 CAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTCAACGGTATGGCTTGGAT 540
QY 541 ATGATGATGAACCTGGTGGCTCAACAGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCTC 600
Db 541 ATGATGATGAACCTGGTGGCTCAACAGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCTC 600
QY 601 TAA 603
Db 601 CAA 603

RESULT 8
AX685022
LOCUS AX685022 723 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 21 from Patent WO0205548.
ACCESSION AX685022
VERSION AX685022.1 GI:29371427
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
          Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, P. and Buyse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
          therapeutic use
JOURNAL Patent: WO 0205548-A 21 18-JUL-2002;
INNOCENTICS N.V. (BE)
FEATURES
    source
        1..723
            /organism="Hepatitis C virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:11103"
        1..720
            /notes="unnamed protein product"
            /codon_start=1
            /protein_id="CAD86525.1"
            /db_xref="GI:29371428"
            /translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLDGG
            VNYATGNLPCGSPSIFELLALLSCLTVPASAYEVNRYVMHVTNDCSNSIYVEAADM
            IMHTPCVPCVRENSSRCWVALTPTLAARNASVPTTIRRHVDSOLFISPRRHETV
            QDCNCSIYPGHITGHMAMDMNWSPTTALVVSQLLRIPQAVVDMVAGHVGVLGAGL
            AYSVMGVNKAUULVLMFLAP"
    CDS
```


REFERENCE 1 (bases 1 to 636)
AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 27 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..636
BASE COUNT 119 a 203 c 174 g 140 t
ORIGIN /organism="unknown"
Query Match 98.6%; Score 597.4; DB 6; Length 636;
Best Local Similarity 99.8%; Pred. No. 1e-121; Indels 0; Gaps 0;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTCAATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGTCAATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGGTCGGGTT 120
Db 61 CCGCTCGTGGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGGTCGGGTT 120
QY 121 CTGGAGGACGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
Db 121 CTGGAGGACGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGTGACCGTTTCCAGCTTTCGGCTTATGAAGTGCACACGTTG 240
Db 181 CTCTTGGCTTTGCTGCTGCTGTGACCGTTTCCAGCTTTCGGCTTATGAAGTGCACACGTTG 240
QY 241 TCGGGATGATACATGCTACAGAACGACTGCTCAAGCAATGCTGATGAGGACGCG 300
Db 241 TCGGGATGATACATGCTACAGAACGACTGCTCAAGCAATGCTGATGAGGACGCG 300
QY 421 ATACGACGCGATCGAATTCGAGCTGTTCAACCATCTCGCCTCGCGGATGAGACGGTG 480
Db 421 ATACGACGCGATCGAATTCGAGCTGTTCAACCATCTCGCCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAATGCTCGCTACACGCGCCCTGCTGATTCGAGCTGCTCCGGATCCT 599
Db 541 ATGATGATGAATGCTCGCTACACGCGCCCTGCTGATTCGAGCTGCTCCGGATCCT 599
RESULT 11
AX452776
LOCUS AX452776 636 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 27 from Patent EP1211315.
ACCESSION AX452776
VERSION AX452776.1 GI:21712461
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 27 05-JUN-2002;
FEATURES Innogenetics N.V. (BE)
Location/Qualifiers

source 1..636
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
CDS 1..633
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38075.1"
/db_xref="GI:21712462"
/translation="MLGKVIDLTTCGFADLVGIVPLVGAPLGGAAALAHGVRLBDC
VNVATNLPCCSFSIFLLALISCLTVPASAVEVNVGVMVHTNDCSNSIVYEADM
INHTPCVPCVRENSSRCWVALFTPLAARNASVPTTIRRHVDSOLFISPRRHETV
QDCNSIYFEHITGHRMWMNMWNSPTTALVVSQLLRIVIEGRHHHHH"
mat_peptide 1..630
/product="unnamed"
BASE COUNT 119 a 203 c 174 g 140 t
ORIGIN
Query Match 98.6%; Score 597.4; DB 6; Length 636;
Best Local Similarity 99.8%; Pred. No. 1e-121; Indels 0; Gaps 0;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTCAATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGTCAATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGGTCGGGTT 120
Db 61 CCGCTCGTGGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGGTCGGGTT 120
QY 121 CTGGAGGACGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
Db 121 CTGGAGGACGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGTGACCGTTTCCAGCTTTCGGCTTATGAAGTGCACACGTTG 240
Db 181 CTCTTGGCTTTGCTGCTGCTGTGACCGTTTCCAGCTTTCGGCTTATGAAGTGCACACGTTG 240
QY 241 TCGGGATGATACATGCTACAGAACGACTGCTCAAGCAATGCTGATGAGGACGCG 300
Db 241 TCGGGATGATACATGCTACAGAACGACTGCTCAAGCAATGCTGATGAGGACGCG 300
QY 301 GACATGATCATGCACACCCCGGTCGCTGCGCTTCCGCTCGCGGATGAGACGGTG 360
Db 301 GACATGATCATGCACACCCCGGTCGCTGCGCTTCCGCTCGCGGATGAGACGGTG 360
QY 361 TGCTGGTAGCGCTACACCCCGGTCGCTGCGCTTCCGCTCGCGGATGAGACGGTG 420
Db 361 TGCTGGTAGCGCTACACCCCGGTCGCTGCGCTTCCGCTCGCGGATGAGACGGTG 420
QY 421 ATACGACGCGATCGAATTCGAGCTGTTCAACCATCTCGCCTCGCGGATGAGACGGTG 480
Db 421 ATACGACGCGATCGAATTCGAGCTGTTCAACCATCTCGCCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAATGCTCGCTACACGCGCCCTGCTGATTCGAGCTGCTCCGGATCCT 599
Db 541 ATGATGATGAATGCTCGCTACACGCGCCCTGCTGATTCGAGCTGCTCCGGATCCT 599
RESULT 12
AX685028
LOCUS AX685028 636 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 27 from Patent WO0205548.
ACCESSION AX685028
VERSION AX685028.1 GI:29371433
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

Mon Dec 22 13:28:41 2003

LOCUS	A48685	Sequence 23 from Patent WO9604385.	561 bp	DNA	linear	PAT 07-MAR-1997	
DEFINITION	A48685						
ACCESSION	A48685						
VERSION	A48685.1	GI:2302398					
KEYWORDS							
SOURCE	unidentified						
ORGANISM	unidentified						
REFERENCE	1 (bases 1 to 561)						
AUTHORS	Maertens,G., Bosman,F., De,M.G. and Buyse,M.						
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE						
JOURNAL	PATENT: WO 9604385-A 23 15-FEB-1996;						
COMMENT	INNOGENETICS NV (BE)						
	Other publication CA 2172273 960215						
	Other publication AU 3382495 960304.						
FEATURES	Location/Qualifiers						
source	1..561	/organism="unidentified"					
		/mol_type="genomic DNA"					
	1..558	/db_xref="taxon:32644"					
CDS		/note="unnamed protein product"					
	1..555	/codon_start=1					
		/protein_id="CAA03135.1"					
		/db_xref="GI:2302398"					
		/translation="MLGKVIDTLTCGFADLVGYIPLVGAFLGGAARALAHGVRLDGG VNYATGNLPGCSFSIFLLALLSLCTVPASAYEVNVMYHVTNDCNSSIVYEAADM IMHTPGCVPCVRENNSSRCWVALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV QDCNCSYFGHITGHRMADMMNMW"					
	1..555	/product="unnamed"					
mat_peptide	103 a	176 c	155 g	127 t			
BASE COUNT	103 a	176 c	155 g	127 t			
ORIGIN							
Query Match	91.7%;	Score 556;	DB 6;	Length 561;			
Best Local Similarity	100.0%;	Pred. No. 1.4e-112;					
Matches	556;	Conservative	0;	Mismatches	0;	Gaps	0;
QY	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATT	60				
Db	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATT	60				
QY	61	CGCTCGTCGGGCCCCCTAGGGGGCGCTGCCAGGCGCTTCCGCGCATGGCGTCCGGGT	120				
Db	61	CGCTCGTCGGGCCCCCTAGGGGGCGCTGCCAGGCGCTTCCGCGCATGGCGTCCGGGT	120				
QY	121	CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTCCTTTCTCTATCTTC	180				
Db	121	CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTCCTTTCTCTATCTTC	180				
QY	181	CTCTTGGCTTTCGCTGCTGACCGTTCAGCTTCCAGCTTCCAGCTTATGAAGTCGCAACGTG	240				
Db	181	CTCTTGGCTTTCGCTGCTGACCGTTCAGCTTCCAGCTTCCAGCTTATGAAGTCGCAACGTG	240				
QY	241	TCCGGGATGACCATGTACGAACTGCTCAACTCAAGCACTCAAGCACTGATGAGGACGG	300				
Db	241	TCCGGGATGACCATGTACGAACTGCTCAACTCAAGCACTCAAGCACTGATGAGGACGG	300				
QY	301	GACATGATCATGCACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCCGC	360				
Db	301	GACATGATCATGCACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCCGC	360				
QY	361	TGCTGGGTAGCGCTCACCCACGCTCGAGCTAGGAACGCCAGCGTCCCAACACGACA	420				
Db	361	TGCTGGGTAGCGCTCACCCACGCTCGAGCTAGGAACGCCAGCGTCCCAACACGACA	420				
QY	421	ATACGACGCCACGTCGATTCACCTTCCAGCTTCCAGCTTCGCGCTCGCGGATGAGACGGTG	480				
Db	421	ATACGACGCCACGTCGATTCACCTTCCAGCTTCCAGCTTCGCGCTCGCGGATGAGACGGTG	480				
QY	481	CAGGACTGCAATTCGCTCAACCGGCGCTTGGTGTATCGAGCTGCTCCGGATCGT	540				
Db	481	CAGGACTGCAATTCGCTCAACCGGCGCTTGGTGTATCGAGCTGCTCCGGATCGT	540				
QY	541	ATGATGATGAACTGGTCCGCTTACAAACGGGCGCTTGGTGTATCGAGCTGCTCCGGATCGT	599				
Db	541	ATGATGATGAACTGGTCCGCTTACAAACGGGCGCTTGGTGTATCGAGCTGCTCCGGATCGT	599				
RESULT 13							
A48685							

Db 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

QY 541 ATGATGATGAACCTGGT 556

Db 541 ATGATGATGAACCTGGT 556

RESULT 14

LOCUS AR157337

DEFINITION Sequence 23 from patent US 6245503.

ACCESSION AR157337

VERSION AR157337.1 GI:16218270

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 561)

AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.

TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use

JOURNAL Patent: US 6245503-A 23 12-JUN-2001;

FEATURES Location/Qualifiers

source 1..561

BASE COUNT 103 a 176 c 155 g 127 t

ORIGIN

Query Match 91.7%; Score 556; DB 6; Length 561;

Best Local Similarity 100.0%; Pred. No. 1.4e-112;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60

Db 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60

QY 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGTCCAGGGCCCTGCGCATGCGTCCGGTT 120

Db 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGTCCAGGGCCCTGCGCATGCGTCCGGTT 120

QY 121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGCTGTGACCGGTTCCAGGTTTCCAGCTTATGAAGTGCACAACGTTG 240

Db 181 CTCTTGGCTTTGCTGCTGTGACCGGTTCCAGGTTTCCAGCTTATGAAGTGCACAACGTTG 240

QY 241 TCCGGGATGTACATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGG 300

Db 241 TCCGGGATGTACATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360

Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGTCAACCCCGGCTAGGAAACGCGGTAGGAAACGCGCTCCCGACGACA 420

Db 361 TGCTGGGTAGCGTCAACCCCGGCTAGGAAACGCGGTAGGAAACGCGCTCCCGACGACA 420

QY 421 ATACGAGCGCACTGATTCGAGCTGTTCAACCATCTCGCTCCCGGATGAGACGGTG 480

Db 421 ATACGAGCGCACTGATTCGAGCTGTTCAACCATCTCGCTCCCGGATGAGACGGTG 480

QY 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

Db 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

QY 541 ATGATGATGAACCTGGT 556

Db 541 ATGATGATGAACCTGGT 556

RESULT 15

AX452772

LOCUS AX452772

DEFINITION Sequence 23 from Patent EP1211315.

ACCESSION AX452772

VERSION AX452772.1 GI:21712457

KEYWORDS Hepatitis C virus

SOURCE Hepatitis C virus

ORGANISM Hepatitis C virus

REFERENCE 1

AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.

TITLE Recombinant vectors for producing hcv envelope proteins

JOURNAL Patent: EP 1211315-A 23 05-JUN-2002;

FEATURES Location/Qualifiers

source 1..561

BASE COUNT 103 a 176 c 155 g 127 t

ORIGIN

Query Match 91.7%; Score 556; DB 6; Length 561;

Best Local Similarity 100.0%; Pred. No. 1.4e-112;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60

Db 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60

QY 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGTCCAGGGCCCTGCGCATGCGTCCGGTT 120

Db 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGTCCAGGGCCCTGCGCATGCGTCCGGTT 120

QY 121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGCTGTGACCGGTTCCAGGTTTCCAGCTTATGAAGTGCACAACGTTG 240

Db 181 CTCTTGGCTTTGCTGCTGTGACCGGTTCCAGGTTTCCAGCTTATGAAGTGCACAACGTTG 240

QY 241 TCCGGGATGTACATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGG 300

Db 241 TCCGGGATGTACATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360

Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGTCAACCCCGGCTAGGAAACGCGGTAGGAAACGCGCTCCCGACGACA 420

Db 361 TGCTGGGTAGCGTCAACCCCGGCTAGGAAACGCGGTAGGAAACGCGCTCCCGACGACA 420

QY 421 ATACGAGCGCACTGATTCGAGCTGTTCAACCATCTCGCTCCCGGATGAGACGGTG 480

Db 421 ATACGAGCGCACTGATTCGAGCTGTTCAACCATCTCGCTCCCGGATGAGACGGTG 480

QY 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

Db 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

us-09-899-303a-25.rge

Mon Dec 22 13:28:41 2003

Db 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTACCGTATGGCTTGGAT 540
Qy 541 ATGATGATGAACTGGT 556
Db 541 ATGATGATGAACTGGT 556

Search completed: December 20, 2003, 02:02:03
Job time : 2295.49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 167.534 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-25
Perfect score: 606
Sequence: 1 AGTTCGGTAAGTCATCGA.....TGCTCGGATCCTCTAATAG 606

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	606	100.0	606	17	AAT12963 HCV E1 construct H
2	606	100.0	606	24	AAL48927 Hepatitis C virus
3	598.2	98.7	723	17	AAT12961 HCV E1 construct H
4	598.2	98.7	723	24	AAL48925 Hepatitis C virus
5	597.4	98.6	636	17	AAT12964 HCV E1 construct H
6	597.4	98.6	636	24	AAL48928 Hepatitis C virus
7	556	91.7	561	17	AAT12962 HCV E1 construct H
8	556	91.7	561	24	AAL48926 Hepatitis C virus

9	516.2	85.2	795	17	AAT12705 HCV E1 construct H
10	516.2	85.2	795	24	AAL48914 Hepatitis C virus
11	513.2	84.7	2082	24	AAL48939 Hepatitis C virus
12	513.2	84.7	2086	17	AAT12973 HCV E1 construct H
13	513.2	84.7	2433	17	AAT12974 HCV E1 construct H
14	502.2	82.9	2434	24	AAL48940 Hepatitis C virus
15	456.4	75.3	633	17	AAT12706 HCV E1 construct H
16	456.4	75.3	633	24	AAL48915 Hepatitis C virus
17	453	74.8	673	19	ABA03491 Cuticle protein 1
18	452.4	74.7	2187	19	ABA03491 NANBHV genomic fra
19	452.4	74.7	2540	15	AQ043889 Hepatitis C virus
20	452.4	74.7	2540	15	AQ063753 Hepatitis C virus
21	452.4	74.7	9605	24	ABK91411 Hepatitis C virus
22	452.4	74.7	9605	24	ABK91424 Hepatitis C virus
23	452.4	74.7	9605	24	ABK91425 Hepatitis C virus
24	452.4	74.7	9605	24	ABK91426 Hepatitis C virus
25	452.4	74.7	9605	24	ABK91428 Hepatitis C virus
26	452.4	74.7	9605	24	ABK91429 Hepatitis C virus
27	452.4	74.7	9605	24	ABK91430 Hepatitis C virus
28	452.4	74.7	9605	24	ABK91431 Hepatitis C virus
29	452.4	74.7	9605	24	ABK91432 Hepatitis C virus
30	452.4	74.7	9605	24	ABK91433 Hepatitis C virus
31	452.4	74.7	9605	24	ABK91434 Hepatitis C virus
32	452.4	74.7	9608	24	ABK91427 Hepatitis C virus
33	452.4	74.7	11062	24	ABK91427 Hepatitis C virus
34	452.4	74.7	11076	21	AA98965 Hepatitis C virus
35	451.6	74.5	636	17	AAT12709 HCV E1 construct H
36	451.6	74.5	636	24	AAL48918 Hepatitis C virus
37	450.8	74.4	1562	19	AAV60672 Fragment #5 isolat
38	450.8	74.4	1880	13	AAQ24467 NANBHV hepatitis vir
39	450.8	74.4	1953	25	AAU55222 Plasmid pIDKE2 DNA
40	450.8	74.4	2829	19	AAV60673 Fragment #6 isolat
41	449.2	74.1	1251	13	AAQ26981 HCV gene 1. Hepat
42	449.2	74.1	2540	13	AAQ26981 Hepatitis C virus
43	449.2	74.1	3360	17	AAT03677 Hepatitis C genome
44	449.2	74.1	3461	15	AAQ64068 Non-A, non-B hepat
45	449.2	74.1	3461	16	AAQ30386 5'UTR/CORE/ENV/NS1

ALIGNMENTS

RESULT 1
AAT12963
ID AAT12963 standard; DNA; 606 BP.

AC AAT12963;

DT 24-SEP-1996 (first entry)

XX HCV E1 construct HCC139.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera; ss.

OS Hepatitis C virus.

PN WO9604385-A2.

XX 15-FEB-1996.

PF 31-JUL-1995; 95WO-EP03031.

PR 29-JUL-1994; 94EP-0870132.

XX (INNO-) INNOGENETICS NV.

PI Bosman F, Buyse M, De Martynoff G, Maertens G;

DR WPI; 1996-129401/13.

PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

PT proteins - in presence of di:sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.

XX Sequence 606 BP; 109 A; 193 C; 167 G; 137 T; 0 other;

Query Match 100.0%; Score 606; DB 17; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.2e-152; Indels 0; Gaps 0;
Matches 606; Conservative 0; Mismatches 0;
QY 1 ATGTTGGTAAAGTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
QY 61 CGCTCGTCTGGCGCCCTTGGGGGGCGCTGCCAGGCGCTGGCGATGCGTCCGGTT 120
DB 61 CGCTCGTCTGGCGCCCTTGGGGGGCGCTGCCAGGCGCTGGCGATGCGTCCGGTT 120
QY 121 CTGAGGACGGGTGAATATGCAACAGGAATTTGCCGGTTCCTTCTCTATCTTC 180
DB 121 CTGAGGACGGGTGAATATGCAACAGGAATTTGCCGGTTCCTTCTCTATCTTC 180
QY 181 CTTCTGGCTTTGCTGCTCTGCTGCTGACCGTTCAGCTTCGGTTATGAGTGGCAACGTG 240
DB 181 CTTCTGGCTTTGCTGCTCTGCTGCTGACCGTTCAGCTTCGGTTATGAGTGGCAACGTG 240
QY 241 TCCGGATGTACATGTACGAAACGACTGCTCCAACTCAAGCTTGTGTATGAGGACGCG 300
DB 241 TCCGGATGTACATGTACGAAACGACTGCTCCAACTCAAGCTTGTGTATGAGGACGCG 300
QY 301 GACATGATCATGACACACCCCGGCTGCGTCCGCTCGGTTTCGGGAGAACAACTTCCGCG 360
DB 301 GACATGATCATGACACACCCCGGCTGCGTCCGCTCGGTTTCGGGAGAACAACTTCCGCG 360
QY 361 TGCTGGGTAGCGTCAACCCCGAGCTTCGAGTGAAGACCGCAGCGTCCACACGACA 420
DB 361 TGCTGGGTAGCGTCAACCCCGAGCTTCGAGTGAAGACCGCAGCGTCCACACGACA 420
QY 421 ATAGAGCCGACGTCGATTTCCAGCTGTTTCCACCTCTCGCTCGCGGATGAGACGGTG 480
DB 421 ATAGAGCCGACGTCGATTTCCAGCTGTTTCCACCTCTCGCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATTTATCCCGGCCAATTAACGGGTCAACGATGCTTGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATTTATCCCGGCCAATTAACGGGTCAACGATGCTTGGAT 540
QY 541 ATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 ATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 TAATAG 606
DB 601 TAATAG 606

DB 601 TAATAG 606

RESULT 2

AA148927
ID AAL48927 standard; DNA; 606 BP.

XX AAL48927;
AC AAL48927;
XX 24-OCT-2002 (first entry)
DT Hepatitis C virus clone HCC139 E1 protein coding sequence.

DE Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.

XX Hepatitis C virus.

XX WO200255548-A2.

XX 18-JUL-2002.

XX 11-JAN-2002; 2002WO-EP00219.

XX 11-JAN-2001; 2001US-260699P.

XX 30-AUG-2001; 2001US-315769P.

XX (INNO-) INNOGENETICS NV.

XX Maertens G, Bosman F, Buyse M;

XX WPI: 2002-599657/64.

XX P-PSDB; AAO18668.

XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection

XX Example 2; Page 177-178; 243pp; English.

XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV) specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.

XX Sequence 606 BP; 109 A; 193 C; 167 G; 137 T; 0 other;

Query Match 100.0%; Score 606; DB 24; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.2e-152; Indels 0; Gaps 0;
Matches 606; Conservative 0; Mismatches 0;

QY 1 ATGTTGGTAAAGTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
QY 61 CGCTCGTCTGGCGCCCTTGGGGGGCGCTGCCAGGCGCTGGCGATGCGTCCGGTT 120
DB 61 CGCTCGTCTGGCGCCCTTGGGGGGCGCTGCCAGGCGCTGGCGATGCGTCCGGTT 120
QY 121 CTGAGGACGGGTGAATATGCAACAGGAATTTGCCGGTTCCTTCTCTATCTTC 180
DB 121 CTGAGGACGGGTGAATATGCAACAGGAATTTGCCGGTTCCTTCTCTATCTTC 180
QY 181 CTCTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

XX Hepatitis C virus clone HCV137 E1 protein coding sequence.
 XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 KW virucide; immunostimulant; vaccine; ds.
 XX Hepatitis C virus.
 OS
 XX WO200255548-A2.
 PN
 XX 18-JUL-2002.
 PD
 XX 11-JAN-2002; 2002WO-EP00219.
 PF
 XX 11-JAN-2001; 2001US-260699P.
 PR
 XX 30-AUG-2001; 2001US-315768P.
 PR
 XX (INNO-) INNOGENETICS NV.
 PA
 XX Maertens G, Bosman F, Buyse M;
 PI
 XX WPI; 2002-599657/64.
 DR
 XX P-PSDB; AAO18666.
 DR
 XX New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans
 PT from HCV infection -
 XX
 XX
 XX Example 2; Page 173-174; 243pp; English.
 PS
 XX The present invention relates to new therapeutic vaccine compositions for
 XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.
 XX
 XX Sequence 723 BP; 126 A; 220 C; 208 G; 169 T; 0 other;
 SQ
 Query Match 98.7%; Score 598.2; DB 24; Length 723;
 Best Local Similarity 99.5%; Pred. No. 1.6e-150;
 Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 1 ATGTTGGTAAGGTATCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
 1 ATGTTGGTAAGGTATCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
 61 CGGCTCGTGGCGCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTT 120
 61 CGGCTCGTGGCGCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTT 120
 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCGCGTGTCTCTTCTATCTTC 180
 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCGCGTGTCTCTTCTATCTTC 180
 181 CTCTTGGCTTGTGTCCTGTGACCGTTCAGCTTCGCTTGAAGTGGCAACGTG 240
 181 CTCTTGGCTTGTGTCCTGTGACCGTTCAGCTTCGCTTGAAGTGGCAACGTG 240
 241 TCCGGATGTACCATGTACGAACTGCTCAACTCAAGCATTTGTGTATGAGCAGCG 300
 241 TCCGGATGTACCATGTACGAACTGCTCAACTCAAGCATTTGTGTATGAGCAGCG 300
 301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360
 301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCCACCGACA 420
 Db 361 TGCTGGGTAGCGCTACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCCACCGACA 420
 QY 421 ATACGAGCCACGCTCGATTCCCGACGCTGTTCCACCATCTCGCTCCCGGATGAGCGGTG 480
 Db 421 ATACGAGCCACGCTCGATTCCCGACGCTGTTCCACCATCTCGCTCCCGGATGAGCGGTG 480
 QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACGCTATGGCTTGGAT 540
 Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACGCTATGGCTTGGAT 540
 QY 541 ATGATGATGAAGTGGTCCGCTACAGCGCCCTGGTGTATCGAGCTGCTCCGGATCTC 600
 Db 541 ATGATGATGAAGTGGTCCGCTACAGCGCCCTGGTGTATCGAGCTGCTCCGGATCTC 600
 QY 601 TAA 603
 Db 601 CAA 603
 RESULT 5
 AAT12964
 ID AAT12964 standard; DNA; 636 BP.
 XX
 AC AAT12964;
 XX
 DT 24-SEP-1996 (first entry)
 XX
 DE HCV E1 construct HCC140.
 XX
 KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.
 XX
 OS Hepatitis C virus.
 XX
 PN WO9604385-A2.
 XX
 PD 15-FEB-1996.
 XX
 PF 31-JUL-1995; 95WO-EP03031.
 XX
 PR 29-JUL-1994; 94EP-0870132.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 XX Bosman F, Buyse M, De Martynoff G, Maertens G;
 XX WPI; 1996-129401/13.
 DR
 XX
 PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 PT proteins - in presence of disulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX
 PS Claim 23; Fig 21; 146pp; English.
 XX
 CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a
 CC disulphide bond cleavage agent, after lysis of recombinant host cells.
 CC The constructs containing the purified HCV envelope proteins can be used
 CC for vaccinating humans against HCV, for in vitro detection of HCV
 CC antibodies in a sample, and in a serotyping assay for detecting one or
 CC more serological types of HCV present in a biological sample. The
 CC constructs can also be immobilised on a solid substrate and incorporated
 CC into a reversed phase hybridisation assay for determining the presence or
 CC the genotype of HCV. The new purification method preserves the

CC	conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.
XX	
SQ	Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;
	Query Match 98.6%; Score 597.4; DB 17; Length 636;
	Best Local Similarity 99.8%; Pred. No. 2.4e-150;
	Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB	1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY	61 CCGCTCGTCCGGCGCCCCCTTAGGGGGCGCTGCCAGAGGCCCTTGGCGGCATGGCGTCCGGGTT 120
DB	61 CCGCTCGTCCGGCGCCCCCTTAGGGGGCGCTGCCAGAGGCCCTTGGCGGCATGGCGTCCGGGTT 120
QY	121 CTGAGAGACGGCGTGAATCTATGCAACAGGGAATTTGGCCGGTTCCTTTCTTCTATCTTC 180
DB	121 CTGAGAGACGGCGTGAATCTATGCAACAGGGAATTTGGCCGGTTCCTTTCTTCTATCTTC 180
QY	181 CTCTTGGCTTTGCTGCTCTGTCACCGTTCCAGCTTCGGCTTATGAAGTCGCAACGTG 240
DB	181 CTCTTGGCTTTGCTGCTCTGTCACCGTTCCAGCTTCGGCTTATGAAGTCGCAACGTG 240
QY	241 TCCGGGATGTACCATGTTCAGCAACGACTGCTCCAACTCAAGCATTTGTATGAGGCAGCG 300
DB	241 TCCGGGATGTACCATGTTCAGCAACGACTGCTCCAACTCAAGCATTTGTATGAGGCAGCG 300
QY	301 GACATGATCATGCACACCCCGGGTGGTGCCTTCGGTTCCGGAGAACAACTCTTCCCGC 360
DB	301 GACATGATCATGCACACCCCGGGTGGTGCCTTCGGTTCCGGAGAACAACTCTTCCCGC 360
QY	361 TGCTGGGTAGGCTCACCCCGACGCTGCAGCTAGGAACGCCAGCGTCCCAACGACACA 420
DB	361 TGCTGGGTAGGCTCACCCCGACGCTGCAGCTAGGAACGCCAGCGTCCCAACGACACA 420
QY	421 ATACGACGCCACGTCGATTCACAGCTGTTACACCATCTCGCCTCGCCGCGCATGAGACGGTG 480
DB	421 ATACGACGCCACGTCGATTCACAGCTGTTACACCATCTCGCCTCGCCGCGCATGAGACGGTG 480
QY	481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCCGTTATGGCTTGGGAT 540
DB	481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCCGTTATGGCTTGGGAT 540
QY	541 ATGATGATGAACCTGGTGCCTTACAAAGCCCTCTGTGTATCGAGCTGCTCCGATCCT 599
DB	541 ATGATGATGAACCTGGTGCCTTACAAAGCCCTCTGTGTATCGAGCTGCTCCGATCCT 599
RESULT 6	
AAI48928	standard; DNA; 636 BP.
ID	AAI48928
XX	
AC	AAI48928;
XX	
DT	24-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus clone HCC140 E1 protein coding sequence.
XX	
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW	virucide; immunostimulant; vaccine; ds.
XX	
OS	Hepatitis C virus.
XX	
PN	W0200255548-A2.
XX	
PD	18-JUL-2002.
XX	
PF	11-JAN-2002; 2002WO-EP00219.
XX	


```

Db      541 ATGATGATGAAGTGGTTCGCTTACAAAGCGCCCTGGTGGTATCGAGCTGCTCCGGATCGT 599
|||||
RESULT 7
AAAT12962
ID      AAT12962 standard; DNA; 561 BP.
XX
AC      AAT12962;
XX
DT      24-SEP-1996 (first entry)
XX
DE      HCV E1 construct HCC138.
XX
KW      HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW      serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW      ss.
XX
OS      Hepatitis C virus.
XX
PN      WO9604385-A2.
XX
PD      15-FEB-1996.
XX
PF      31-JUL-1995; 95WO-EP03031.
XX
PR      29-JUL-1994; 94EP-0870132.
XX
PA      (INNO-) INNOGENETICS NV.
XX
PI      Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
DR      WPI; 1996-129401/13.
XX
PT      Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT      proteins - in presence of disulphide bond cleavage agent, to
PT      produce proteins suitable for direct use in vaccines or diagnostic
PT      assays of HCV
XX
PS      Claim 23; Fig 21; 146pp; English.
XX
XX      AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX      and E2 protein coding sequence constructs. These sequences are included
XX      in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX      The recombinant proteins can then be isolated using a method of the
XX      invention. In the method, the envelope proteins are purified by
XX      carrying out a disulphide bond cleavage, or a reduction step with a
XX      disulphide bond cleavage agent, after lysis of recombinant host cells.
XX      The constructs containing the purified HCV envelope proteins can be used
XX      for vaccinating humans against HCV, for in vitro detection of HCV
XX      antibodies in a sample, and in a serotyping assay for detecting one or
XX      more serological types of HCV present in a biological sample. The
XX      constructs can also be immobilised on a solid substrate and incorporated
XX      into a reversed phase hybridisation assay for determining the presence or
XX      the genotype of HCV. The new purification method preserves the
XX      conformation of the recombinantly expressed E1, E2 and E1/E2, and
XX      eliminates contaminating proteins. Antigens isolated using this method
XX      are more reactive with human sera than those isolated by known
XX      techniques.
XX
SQ      Sequence 561 BP; 103 A; 176 C; 155 G; 127 T; 0 other;
Query Match      91.7%; Score 556; DB 17; Length 561;
Best Local Similarity 100.0%; Pred. No. 2.9e-139;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 ATGTTGGTAAAGTTCATCGATACCTTACATGCGGCTTCCCGACCTCTGTGGGTACATT 60
Db      1 ATGTTGGTAAAGTTCATCGATACCTTACATGCGGCTTCCCGACCTCTGTGGGTACATT 60
Qy      61 CCGCTCGTGGCGCCCCCTAGGGGGCGTCCAGGGCCCTGGCGGATGCGTCCGGGT 120
Db      61 CCGCTCGTGGCGCCCCCTAGGGGGCGTCCAGGGCCCTGGCGGATGCGTCCGGGT 120

```

```

Qy      121 CTGAGGACGGCGTGAATATGCAACAGGAATTTGCCGGTGGCTCTTTCTATCTTC 180
Db      121 CTGAGGACGGCGTGAATATGCAACAGGAATTTGCCGGTGGCTCTTTCTATCTTC 180
Qy      181 CTCTTGGCTTTGCTGTCCTGTTCTGACCGGTTCCAGCTTCCGGTTATGAAGTGGCAACGTG 240
Db      181 CTCTTGGCTTTGCTGTCCTGTTCTGACCGGTTCCAGCTTCCGGTTATGAAGTGGCAACGTG 240
Qy      241 TCCGGGATGTACCATGTCTCAGAACGACTGCTCAACTCAAGCATTTGTGTATGAGGACGG 300
Db      241 TCCGGGATGTACCATGTCTCAGAACGACTGCTCAACTCAAGCATTTGTGTATGAGGACGG 300
Qy      301 GACATGATCATGACACACCCCGGTCCTGCTCGGAGAACCAACTCTTCCCGC 360
Db      301 GACATGATCATGACACACCCCGGTCCTGCTCGGAGAACCAACTCTTCCCGC 360
Qy      361 TGCTGGGTAGCGTTCACCCCGCTCGAGCTAGGAAAGCCAGCGTCCCGACCGACGACA 420
Db      361 TGCTGGGTAGCGTTCACCCCGCTCGAGCTAGGAAAGCCAGCGTCCCGACCGACGACA 420
Qy      421 ATACGACGCGACGTCGATTCGAGCTGTTCCACCATCTCGCCTCGCGGATGAGACGGTG 480
Db      421 ATACGACGCGACGTCGATTCGAGCTGTTCCACCATCTCGCCTCGCGGATGAGACGGTG 480
Qy      481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTCAACGATATGGCTTGGGAT 540
Db      481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACCGGTCAACGATATGGCTTGGGAT 540
Qy      541 ATGATGATGAAGTGGT 556
Db      541 ATGATGATGAAGTGGT 556
RESULT 8
AAL48926
ID      AAL48926 standard; DNA; 561 BP.
XX
AC      AAL48926;
XX
DT      24-OCT-2002 (first entry)
XX
DE      Hepatitis C virus clone HCC138 E1 protein coding sequence.
XX
KW      Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW      virucide; immunostimulant; vaccine; ds.
XX
OS      Hepatitis C virus.
XX
PN      WO200255548-A2.
XX
PD      18-JUL-2002.
XX
PF      11-JAN-2002; 2002WO-EP00219.
XX
PR      11-JAN-2001; 2001US-260699P.
XX
PR      30-AUG-2001; 2001US-315788P.
XX
PA      (INNO-) INNOGENETICS NV.
XX
PI      Maertens G, Bosman F, Buyse M;
XX
DR      WPI; 2002-599657/64.
XX
PT      P-PSDB; AAO18667.
XX
PT      New therapeutic vaccine compositions comprising at least one purified
PT      recombinant hepatitis C virus (HCV) single or specific oligomeric
PT      recombinant envelope protein E1 or E2, useful for immunizing humans
PT      from HCV infection.
XX
PS      Example 2; Page 175-176; 243pp; English.
XX
CC      The present invention relates to new therapeutic vaccine compositions for

```


CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.
 XX
 XX Sequence 561 BP; 103 A; 176 C; 155 G; 127 T; 0 other;

Query Match 91.7%; Score 556; DB 24; Length 561;
 Best Local Similarity 100.0%; Pred. No. 2.9e-139;
 Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTTGGGTAAGGTATCGATACCCCTTACATGGGCTTCGCCAGCTCGTGGGTACATT 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1 ATGTTGGGTAAGGTATCGATACCCCTTACATGGGCTTCGCCAGCTCGTGGGTACATT 60
 QY 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGGTGCCAGGGCCCTGCGCATGGCGTCCGGGT 120
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGGTGCCAGGGCCCTGCGCATGGCGTCCGGGT 120
 QY 121 CTGGAGAGCGGTGAACATATGATCAACAGGGAATTTCCCGGTTGCTTTCTATCTTC 180
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 121 CTGGAGAGCGGTGAACATATGATCAACAGGGAATTTCCCGGTTGCTTTCTATCTTC 180
 QY 181 CTCTTGGCTTGTCTGCTGTCTGAGCGTTCAGCTTCGCTTATGAAGTGCACAGTG 240
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 181 CTCTTGGCTTGTCTGCTGTCTGAGCGTTCAGCTTCGCTTATGAAGTGCACAGTG 240
 QY 241 TCCGGATGTACCATGTCAAGACGACTGTCTCAACTCAAGCATTTGTATGAGGAGCG 300
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 241 TCCGGATGTACCATGTCAAGACGACTGTCTCAACTCAAGCATTTGTATGAGGAGCG 300
 QY 301 GACATGATCATGCACACCCCGGGTGCCTGCTGCTTGGGAGAACAACTCTTCCCGC 360
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 301 GACATGATCATGCACACCCCGGGTGCCTGCTGCTTGGGAGAACAACTCTTCCCGC 360
 QY 361 TCGTGGTGTAGCGTCAACCCCGAGCTCGAGCTAGGAAAGCCAGCGTCCCAACAGACA 420
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 361 TCGTGGTGTAGCGTCAACCCCGAGCTCGAGCTAGGAAAGCCAGCGTCCCAACAGACA 420
 QY 421 ATACGAGCGCAGCTCGATTTCCAGCTGTTCAACCATCTCGCTCGCGGATGAGAGCGTG 480
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 421 ATACGAGCGCAGCTCGATTTCCAGCTGTTCAACCATCTCGCTCGCGGATGAGAGCGTG 480
 QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTTGGGAT 540
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTTGGGAT 540
 QY 541 ATGATGATGAACGTGGT 556
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 541 ATGATGATGAACGTGGT 556

RESULT 9
 AAT12705
 ID AAT12705 standard; DNA; 795 BP.
 XX
 AC AAT12705;
 XX
 DT 23-SEP-1996 (first entry)
 DE HCV E1 construct HCC110A.
 XX
 KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 XX 55.

OS Hepatitis C virus.
 XX
 PN WO9604385-A2.
 XX
 PD 15-FEB-1996.
 XX
 PF 31-JUL-1995; 95WO-EP03031.
 XX
 PR 29-JUL-1994; 94BP-0870132.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Bosman F, Buyse M, De Martynoff G, Maertens G;
 XX WPI; 1996-129401/13.
 DR
 XX
 PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 PT proteins - in presence of disulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX
 PS Claim 23; Fig 21; 146pp; English.
 XX
 CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a
 CC disulphide bond cleavage agent, after lysis of recombinant host cells.
 CC The constructs containing the purified HCV envelope proteins can be used
 CC for vaccinating humans against HCV, for in vitro detection of HCV
 CC antibodies in a sample, and in a serotyping assay for detecting one or
 CC more serological types of HCV present in a biological sample. The
 CC constructs can also be immobilised on a solid substrate and incorporated
 CC into a reversed phase hybridisation assay for determining the presence or
 CC the genotype of HCV. The new purification method preserves the
 CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
 CC eliminates contaminating proteins. Antigens isolated using this method
 CC are more reactive with human sera than those isolated by known
 CC techniques.
 XX
 SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match 85.2%; Score 516.2; DB 17; Length 795;
 Best Local Similarity 88.9%; Pred. No. 1.5e-128;
 Matches 600; Conservative 0; Mismatches 3; Indels 72; Gaps 1;
 QY 1 ATGTTGGGTAAGGTATCGATACCCCTTACATGGGCTTCGCCAGCTCGTGGGTACATT 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1 ATGTTGGGTAAGGTATCGATACCCCTTACATGGGCTTCGCCAGCTCGTGGGTACATT 60
 QY 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGGTGCCAGGGCCCTGCGCATGGCGTCCGGGT 120
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGGTGCCAGGGCCCTGCGCATGGCGTCCGGGT 120
 QY 121 CTGGAGAGCGGTGAACATATGATCAACAGGGAATTTCCCGGTTGCTTTCTATCTTC 180
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 121 CTGGAGAGCGGTGAACATATGATCAACAGGGAATTTCCCGGTTGCTTTCTATCTTC 180
 QY 181 CTCTTGGCTTGTCTGCTGTCTGAGCGTTCAGCTTCGCTTATGAAGTGCACAGTG 240
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 181 CTCTTGGCTTGTCTGCTGTCTGAGCGTTCAGCTTCGCTTATGAAGTGCACAGTG 240
 QY 241 TCCGGATGTACCATGTCAAGACGACTGTCTCAACTCAAGCATTTGTATGAGGAGCG 300
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 241 TCCGGATGTACCATGTCAAGACGACTGTCTCAACTCAAGCATTTGTATGAGGAGCG 300
 QY 301 GACATGATCATGCACACCCCGGGTGCCTGCTGCTTGGGAGAACAACTCTTCCCGC 360
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 301 GACATGATCATGCACACCCCGGGTGCCTGCTGCTTGGGAGAACAACTCTTCCCGC 360
 QY 361 TGCTGGTGTAGCGTCAACCCCGAGCTCGAGCTAGGAAAGCCAGCGTCCCAACAGACA 420

CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match 85.2%; Score 516.2; DB 24; Length 795;
Best Local Similarity 88.9%; Pred. No. 1.5e-128;
Matches 600; Conservative 0; Mismatches 3; Indels 72; Gaps 1;

QY 1 ATCTGGGTAAGGTATCATGATACCTTACATCGGCTTCGCGACCTCGTGGGGTACATTT 60
DB 1 ATCTGGGTAAGGTATCATGATACCTTACATCGGCTTCGCGACCTCGTGGGGTACATTT 60
QY 61 CCCTCGTGGCGCCCTCCCTAGGGGCGCTGCGAGGCGCTGCGCATCGCGGTT 120
DB 61 CCGCTCGTGGCGCCCTCCCTAGGGGCGCTGCGAGGCGCTGCGCATCGCGGTT 120
QY 121 CTGAGGACCGCGTGAATATGCAAGGGAATTTGCCGGTTGCTCTTCTATCTTTC 180
DB 121 CTGAGGACCGCGTGAATATGCAAGGGAATTTGCCGGTTGCTCTTCTATCTTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
QY 241 TCCGGGATGTACCATGTACGAAACGACTGCTCAACTCAAGCATTTGTATGAGGCGG 300
DB 241 TCCGGGATGTACCATGTACGAAACGACTGCTCAACTCAAGCATTTGTATGAGGCGG 300
QY 301 GACATGATCATGACACCGCGGTGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACCGCGGTGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGCTCGAGTAGGAAACGCGCTGCCACACGACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGCTCGAGTAGGAAACGCGCTGCCACACGACA 420
QY 421 ATACGACGCGACGTCGAT-----TCCAGCTGTTTACCATCTCGCTCGCGG 468
DB 421 ATACGACGCGACGTCGAT-----TCCAGCTGTTTACCATCTCGCTCGCGG 468
QY 439 -----TCCAGCTGTTTACCATCTCGCTCGCGG 540
DB 439 -----TCCAGCTGTTTACCATCTCGCTCGCGG 540
QY 469 CATGAGACGCTGACGACTGCAATTTCTCAATCTATCCGGCCACATAACGGGTACCGT 528
DB 469 CATGAGACGCTGACGACTGCAATTTCTCAATCTATCCGGCCACATAACGGGTACCGT 528
QY 541 CATGAGACGCTGACGACTGCAATTTCTCAATCTATCCGGCCACATAACGGGTACCGT 600
DB 541 CATGAGACGCTGACGACTGCAATTTCTCAATCTATCCGGCCACATAACGGGTACCGT 600
QY 529 ATGGCTTGGGATATGATGATGAATGCTGCGCTACAAAGCGCCCTGGTGTATCGAGCTG 588
DB 529 ATGGCTTGGGATATGATGATGAATGCTGCGCTACAAAGCGCCCTGGTGTATCGAGCTG 588
QY 601 ATGGCTTGGGATATGATGATGAATGCTGCGCTACAAAGCGCCCTGGTGTATCGAGCTG 660
DB 601 ATGGCTTGGGATATGATGATGAATGCTGCGCTACAAAGCGCCCTGGTGTATCGAGCTG 660
QY 589 CTCGGATCTCTAA 603
DB 589 CTCGGATCTCTAA 603
QY 661 CTCGGATCTCTAA 675
DB 661 CTCGGATCTCTAA 675

RESULT 11

AAAL48939
ID AAL48939 standard; DNA; 2082 BP.

XX
AC AAL48939;
XX
DT 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.
XX
KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX
KW virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
PN WO20025548-A2.

DB 361 TGCTGGGTAGCGCTCACCCCAACGCTCGAGCTAGAAACGACGGTCCCAACACGACA 420
QY 421 ATACGAGCGCACGTCGAT-----TCCAGCTGTTTACCATCTCGCTCGCGG 468
DB 421 ATACGAGCGCACGTCGAT-----TCCAGCTGTTTACCATCTCGCTCGCGG 468
QY 439 -----TCCAGCTGTTTACCATCTCGCTCGCGG 540
DB 439 -----TCCAGCTGTTTACCATCTCGCTCGCGG 540
QY 469 CATGAGACGCTGACGACTGCAATTTCTCAATCTATCCGGCCACATAACGGGTACCGT 528
DB 469 CATGAGACGCTGACGACTGCAATTTCTCAATCTATCCGGCCACATAACGGGTACCGT 528
QY 541 CATGAGACGCTGACGACTGCAATTTCTCAATCTATCCGGCCACATAACGGGTACCGT 600
DB 541 CATGAGACGCTGACGACTGCAATTTCTCAATCTATCCGGCCACATAACGGGTACCGT 600
QY 529 ATGGCTTGGGATATGATGATGAATGCTGCGCTACAAAGCGCCCTGGTGTATCGAGCTG 588
DB 529 ATGGCTTGGGATATGATGATGAATGCTGCGCTACAAAGCGCCCTGGTGTATCGAGCTG 588
QY 601 ATGGCTTGGGATATGATGATGAATGCTGCGCTACAAAGCGCCCTGGTGTATCGAGCTG 660
DB 601 ATGGCTTGGGATATGATGATGAATGCTGCGCTACAAAGCGCCCTGGTGTATCGAGCTG 660
QY 589 CTCGGATCTCTAA 603
DB 589 CTCGGATCTCTAA 603
QY 661 CTCGGATCTCTAA 675
DB 661 CTCGGATCTCTAA 675

RESULT 10

AAAL48914
ID AAL48914 standard; DNA; 795 BP.

XX
AC AAL48914;
XX
DT 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus clone HCC110A E1 protein coding sequence.
XX
DE Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX
KW virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
PN WO20025548-A2.

18-JUL-2002.

11-JAN-2002; 2002WO-EP00219.

11-JAN-2001; 2001US-260699P.

30-AUG-2001; 2001US-315768P.

(INNO-) INNOGENETICS NV.

Maertens G, Bosman F, Buyse M;

WPI; 2002-599657/64.

P-PSDB; AAO18661.

XX
PT New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection.

XX
PS Example 2; Page 161-162; 243pp; English.

XX
CC The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence

XX 18-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-EF00219.
 XX
 PR 11-JAN-2001; 2001US-260699P.
 PR 30-AUG-2001; 2001US-315768P.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Maertens G, Bosman F, Buyse M;
 DR WPI; 2002-599657/64.
 DR P-PSDB; AA018678.
 XX
 PT New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans
 PT from HCV infection
 XX
 PS Example 2; Page 206-209; 243pp; English.
 XX
 CC The present invention relates to new therapeutic vaccine compositions for
 CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope protein selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.
 XX
 SQ Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 other;

Query Match 84.7%; Score 513.2; DB 24; Length 2082;
 Best Local Similarity 88.8%; Pred. No. 1.3e-127;
 Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;

QY 4 TTGGGTAAGTCAATCATACCTTACATCGGCTTCGCGACCTCGTGGGTACATTCGG 63
 DB 4 TTGGGTAAGTCAATCATACCTTACATCGGCTTCGCGACCTCGTGGGTACATTCGG 63
 QY 64 CTCGTGCGGCGCCCTAGGGGCGCTGCGAGGCGCTGCGCATGCGGTCGGGTTCTG 123
 DB 64 CTCGTGCGGCGCCCTAGGGGCGCTGCGAGGCGCTGCGCATGCGGTCGGGTTCTG 123
 QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTTTCTATCTTCCTC 183
 DB 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTTTCTATCTTCCTC 183
 QY 184 TTGGCTTGTCTGCTGCTGACCGTTCAGCTTCGCTTATGAAGTGCGACGTGTC 243
 DB 184 TTGGCTTGTCTGCTGCTGACCGTTCAGCTTCGCTTATGAAGTGCGACGTGTC 243
 QY 244 GGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGGAC 303
 DB 244 GGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGGAC 303
 QY 304 ATGATCATGCACACCCCGGGTGCCTGCTTCGGAGAACAACTCTTCCCGCTGC 363
 DB 304 ATGATCATGCACACCCCGGGTGCCTGCTTCGGAGAACAACTCTTCCCGCTGC 363
 QY 364 TGGGTAGCGTACCCCGAGCTCGAGTAGGAAAGCCAGCGTCCCGACGACAAATA 423
 DB 364 TGGGTAGCGTACCCCGAGCTCGAGTAGGAAAGCCAGCGTCCCGACGACAAATA 423
 QY 424 CGACGCCACGTGCAT-----
 DB 424 CGACGCCACGTGCATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGTAGTGGGG 483

QY 439 -----TCCAGCTGTTCCACATCTCGCTCCCGGCAT 471
 DB 484 GACCTCTCGGATCTGTCTTCCTCGCTCCCGCATTTCCACATCTCGCTCCCGCAT 543
 QY 472 GAGACGGTGCAGGACTGCAATTCATCCCGGCACATAACGGGTACCGGTATG 531
 DB 544 GAGACGGTGCAGGACTGCAATTCATCCCGGCACATAACGGGTACCGGTATG 603
 QY 532 GCTTGGATATGATGATGAATGCTGCTCAACAGCGCCCTGGTGTATCGCAGCTGCTC 591
 DB 604 GCTTGGATATGATGATGAATGCTGCTCAACAGCGCCCTGGTGTATCGCAGCTGCTC 663
 QY 592 CGGATCCTCTAA 603
 DB 664 CGGATCCCAAA 675
 RESULT 12
 AAT12973
 ID AAT12973 standard; DNA; 2086 BP.
 XX
 AC AAT12973;
 XX
 DT 24-SEP-1996 (first entry)
 XX
 DE HCV E1 construct HCC165.
 XX
 KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.
 XX
 OS Hepatitis C virus.
 XX
 PN WO9604385-A2.
 XX
 PD 15-FEB-1996.
 XX
 PF 31-JUL-1995; 95WO-EF03031.
 XX
 PR 29-JUL-1994; 94EP-0870132.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Bosman F, Buyse M, De Martynoff G, Maertens G;
 XX
 DR WPI; 1996-129401/13.
 XX
 PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 PT proteins - in presence of di: sulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX
 PS Claim 23; Fig 21; 146pp; English.
 XX
 CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a
 CC disulphide bond cleavage agent, after lysis of recombinant host cells.
 CC The constructs containing the purified HCV envelope proteins can be used
 CC for vaccinating humans against HCV, for in vitro detection of HCV
 CC antibodies in a sample, and in a serotyping assay for detecting one or
 CC more serological types of HCV present in a biological sample. The
 CC constructs can also be immobilised on a solid substrate and incorporated
 CC into a reversed phase hybridisation assay for determining the presence or
 CC the genotype of HCV. The new purification method preserves the
 CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
 CC eliminates contaminating proteins. Antigens isolated using this method
 CC are more reactive with human sera than those isolated by known
 CC techniques.
 XX

15-FEB-1996.

31-JUL-1995; 95WO-EP03031.

29-JUL-1994; 94EP-0870132.

(INNO-) INNOGENETICS NV.

Bosman P, Buyse M, De Martynoff G, Maertens G; WPI, 1996-129401/13.

Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - in presence of di: sulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

Claim 23; Fig 21; 146pp; English.

AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;

Query Match 84.7%; Score 513.2; DB 17; Length 2433;
Best Local Similarity 88.8%; Pred. No. 1.3e-127;
Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;

QY 4 TTGGGTAAAGTTCATCGATACATCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCGG 63
Db 355 TTGGGTAAAGTTCATCGATACATCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCGG 414
QY 64 CTCGTGGCGCCCCCTAGGGGGCGCTGCCAGAGGCCCTGGCGCATGCGTCCGGGTCTG 123
Db 415 CTCGTGGCGCCCCCTAGGGGGCGCTGCCAGAGGCCCTGGCGCATGCGTCCGGGTCTG 474
QY 124 GAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTCTC 183
Db 475 GAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTCTC 534
QY 184 TTGGCTTTGCTGCTGCTGCTGACCCCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 243
Db 535 TTGGCTTTGCTGCTGCTGCTGACCCGTTCCAGCTTCGGCTTATGAAGTCGCAACGTGTC 594
QY 244 GGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTGTATGAGCGACGGAC 303
Db 595 GGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTGTATGAGCGACGGAC 654
QY 304 ATGATCATGCAACACCCCGGGTGGCTGCGCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 363
Db 655 ATGATCATGCAACACCCCGGGTGGCTGCGCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 714
QY 364 TGGGTAGCGCTCAACCCCGGGTGGCTGCGCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 423
Db 715 TGGGTAGCGCTCAACCCCGGGTGGCTGCGCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 774
QY 424 CGACGCCACGTTCGAT----- 438

Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;

Query Match 84.7%; Score 513.2; DB 17; Length 2086;
Best Local Similarity 88.8%; Pred. No. 1.3e-127;
Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;

QY 4 TTGGGTAAAGTTCATCGATACATCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCGG 63
Db 4 TTGGGTAAAGTTCATCGATACATCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCGG 63
QY 64 CTCGTGGCGCCCCCTAGGGGGCGCTGCCAGAGGCCCTGGCGCATGCGTCCGGGTCTG 123
Db 64 CTCGTGGCGCCCCCTAGGGGGCGCTGCCAGAGGCCCTGGCGCATGCGTCCGGGTCTG 123
QY 124 GAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTCTC 183
Db 124 GAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTCTC 183
QY 184 TTGGCTTTGCTGCTGCTGCTGACCCCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 243
Db 184 TTGGCTTTGCTGCTGCTGCTGACCCGTTCCAGCTTCGGCTTATGAAGTCGCAACGTGTC 243
QY 244 GGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTGTATGAGCGACGGAC 303
Db 244 GGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTGTATGAGCGACGGAC 303
QY 304 ATGATCATGCAACACCCCGGGTGGCTGCGCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 363
Db 304 ATGATCATGCAACACCCCGGGTGGCTGCGCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 363
QY 364 TGGGTAGCGCTCAACCCCGGGTGGCTGCGCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 423
Db 364 TGGGTAGCGCTCAACCCCGGGTGGCTGCGCTTCAGCTTCGGCTTATGAAGTCGCAACGTGTC 423
QY 424 CGACGCCACGTTCGAT----- 438
Db 424 CGACGCCACGTTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGGG 483
QY 439 -----TCCAGCTGTTTCCACCATCTCGCTCGCGGAT 471
Db 484 GACCTCTCGGATCTGCTTCTCGCTCCAGCTGTTCCACCATCTCGCTCGCGGAT 543
QY 472 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGGTACCGGTATG 531
Db 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGGTACCGGTATG 603
QY 532 GCTTGGGATATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
Db 604 GCTTGGGATATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
QY 592 CGGATCTCTTAA 603
Db 664 CGGATCTCTTAA 675

RESULT 13
AAT12974
ID AAT12974 standard; DNA; 2433 BP.

XX AC AAT12974;
XX AC AAT12974;
XX AC AAT12974;
DT 25-SEP-1996 (first entry)
DE HCV E1 construct HCCI66.
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX Hepatitis C virus.
OS
XX W09604385-A2.
PN
XX


```
|||||
775 CGAGCCACGTCGATTTGCTGCTGGGGCGGCTGCTTTCTTCCGCTATGTAAGTGGG 834
439 -----TCCAGCTGTTTACCACATCTCGCCTCGCCGGCAT 471
835 GACCTCTGCGGATCTGTCCTCTCCAGCTGTTTACCACATCTCGCCTCGCCGGCAT 894
472 GAGACGGTGAGGACTGCAATTTGCTCAATATTCGCGGCCACATAACGGGTACCGTATG 531
895 GAGACGGTGAGGACTGCAATTTGCTCAATATTCGCGGCCACATAACGGGTACCGTATG 954
532 GCTTGGGATATGATGTAAGTAACTGCTCGCTTACAAAGCGCCCTGGTATGCGAGCTGCTC 591
955 GCTTGGGATATGATGTAAGTAACTGCTCGCTTACAAAGCGCCCTGGTATGCGAGCTGCTC 1014
592 CGGATCCTCTTAA 603
1015 CGGATCCCAAA 1026

RESULT 14
AAL48940
ID AAL48940 standard; DNA; 2434 BP.
XX
AC AAL48940;
XX
DT 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.
XX
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus..
XX
XX WO200255548-A2.
XX
PD 18-JUL-2002.
XX
PF 11-JAN-2002; 2002WO-EP00219.
XX
PR 11-JAN-2001; 2001US-260699P.
XX
PR 20-AUG-2001; 2001US-315768P.
XX
PA (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Bosman F, Buysse M;
XX
XX WPI; 2002-599657/64.
XX
XX P-PSDB; AAO18679.
XX
XX New therapeutic vaccine compositions comprising at least one purified
XX recombinant hepatitis C virus (HCV) single or specific oligomeric
XX recombinant envelope protein E1 or E2, useful for immunizing humans
XX from HCV infection
XX
XX Example 2; Page 212-215; 243pp; English.
XX
XX The present invention relates to new therapeutic vaccine compositions for
XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
XX composition containing at least one purified recombinant HCV single or
XX specific oligomeric recombinant envelope proteins selected from an E1 and
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
XX useful for inducing HCV-specific antibodies or for immunising humans
XX against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
XX vaccines or therapeutics, in HCV screening and confirmatory antibody
XX tests, for raising antibodies, in the preparation of medicament, and for
XX in vitro monitoring of HCV disease or prognosing the response to
XX treatment of patients suffering from HCV infection. The present sequence
XX is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 other;
```

```
Query Match 82.9%; Score 502.2; DB 24; Length 2434;
Best Local Similarity 88.7%; Pred. No. 1.2e-124;
Matches 597; Conservative 0; Mismatches 3; Indels 73; Gaps 2;

QY 4 TTGGGTAAGTGCATCGATACCCCTTACATGGGCTTCGCGACCTCGTGGGTACATTCGG 63
DB |||||
DB 355 TTGGGTAAGTGCATCGATACCCCTTACATGGGCTTCGCGACCTCGTGGGTACATTCGG 414
QY 64 CTGCTCGGGGCCCCCTTAGGGGCGCTGCGACGGCCCTGGCGCATGGCGTTCGGTTCG 123
DB |||||
DB 415 CTGCTCGGGGCCCCCTTAGGGGCGCTGCGACGGCCCTGGCGCATGGCGTTCGGTTCG 474
QY 124 GAGGACGGCGTGAACTATGATCAACAGGGAATTTCCCGGTTGCTCTTCTATCTTCCTC 183
DB |||||
DB 475 GAGGACGGCGTGAACTATGATCAACAGGGAATTTCCCGGTTGCTCTTCTATCTTCCTC 534
QY 184 TTGGCTTTTGTCTGCTG-TCTGACCGTTCGACCTTCGCGTTCGCGTTCGCGTTCGCGTTC 242
DB |||||
DB 535 TTGGCTTTTGTCTGCTGCTGACCCGTTCCAGCTTCGCGTTCGCGTTCGCGTTCGCGTTC 594
QY 243 CGGGATGTATACCATGTTCACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGA 302
DB |||||
DB 595 CGGGATGTATACCATGTTCACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGA 654
QY 303 CATGATCATGCACACCCCGGGTGGTGGCTGCGTTCGGGAGAAACAATCTTTCCGCTG 362
DB |||||
DB 655 CATGATCATGCACACCCCGGGTGGTGGCTGCGTTCGGGAGAAACAATCTTTCCGCTG 714
QY 363 CTGGGTAGGGCTCACCCCGGCTGCGAGCTAGGACGCGGCTCCCGACGACGCAAT 422
DB |||||
DB 715 CTGGGTAGGGCTCACCCCGGCTGCGAGCTAGGACGCGGCTCCCGACGACGCAAT 774
QY 423 ACGACGCCACGTCGAT-----TTCGGGGGGGCTGCTTTCTGCTATGTAAGTGGG 438
DB |||||
DB 775 ACGACGCCACGTCGATTTGCTTGGGGGGGCTGCTTTCTGCTATGTAAGTGGG 834
QY 439 -----TCCAGCTGTTTACCACATCTCGCCTCGCCGGCA 470
DB |||||
DB 835 GGAACCTCTCGGATCTGCTTCTCGCTCCAGCTGTTTCCACCATCTCGCCTCGCCGGCA 894
QY 471 TGAGACGGTGACGAGTGCATTTGCTCAATCTATCCGGCCACATTAACGGGTACCGTAT 530
DB |||||
DB 895 TGAGACGGTGACGAGTGCATTTGCTCAATCTATCCGGCCACATTAACGGGTACCGTAT 954
QY 531 GCTTGGGATATGATGATGAATGCTCGCTTACAAAGCGCCCTGCTGCTATGCGAGCTGCT 590
DB |||||
DB 955 GCTTGGGATATGATGATGAATGCTCGCTTACAAAGCGCCCTGCTGCTATGCGAGCTGCT 1014
QY 591 CCGGATCCTCTTAA 603
DB |||||
DB 1015 CCGGATCCCAAA 1026

RESULT 15
AAT12706
ID AAT12706 standard; DNA; 633 BP.
XX
XX AAT12706;
XX
XX AC AAT12706;
XX
XX DT 23-SEP-1996 (first entry)
XX
XX DE HCV E1 construct HCC111A.
XX
XX KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX
XX OS Hepatitis C virus.
XX
XX PN WO9604385-A2.
XX
XX PD 15-FEB-1996.
XX
```


us-09-899-303a-25.rng

Mon Dec 22 13:28:41 2003

439 -----TCCAGCTGTTCCACCATCTCGCTCGCCGG 468
Qy
481 GGGGATCTCTGGGATCTCTCTTCCTCGTCTCCAGCTGTTCCACCATCTCGCTCGCCGG 540
Db
469 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT 528
Qy
541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT 600
Db
529 ATGGCTTGGGATATGATGAACTGGT 556
Qy
601 ATGGCTTGGGATATGATGAACTGGT 628
Db
Search completed: December 19, 2003, 18:51:18
Job time : 170.534 secs

PF 31-JUL-1995; 95WO-EP03031.
XX PR
XX PR 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX PA Bosman F, Buyse M, De Martynoff G, Maertens G;
XX PI WPI; 1996-129401/13.
XX DR
XX XX
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di: sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
XX Claim 23; Fig 21; 146pp; English.
XX
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
XX Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;
SQ
Query Match 75.3%; Score 456.4; DB 17; Length 633;
Best Local Similarity 86.8%; Pred. No. 1.5e-112;
Matches 545; Conservative 0; Mismatches 11; Indels 72; Gaps 1;
Qy 1 ATGTTGGGTAGGTATCATGATACCTTACATCGCGCTTCGCGGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAGGTATCATGATACCTTACATCGCGCTTCGCGGACCTCATGGGTACATT 60
Qy 61 CCGCTCGTCGGCGCCCGCCCTAGGGGGCGCTGCGAGGGCCCTGGGCGATGGCGTCCGGTT 120
Db 61 CCGCTCGTCGGCGCCCGCCCTAGGGGGCGCTGCGAGGGCCCTGGGCGATGGCGTCCGGTT 120
Qy 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC 180
Qy 181 CTCTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTG 240
Db 181 CTCTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTG 240
Qy 241 TCCGGGATGATCATGTCACGAACTGCTCAACTCAAGCAATTTGTGTATGAGCAGCG 300
Db 241 TCCGGGATGATCATGTCACGAACTGCTCAACTCAAGCAATTTGTGTATGAGCAGCG 300
Qy 301 GACATGATCATGACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCCCGC 360
Qy 361 TGCTGGGTAGCGCTACCCCGAGCTCGCAGCTAGGAAAGCCAGCGTCCCGCAGACA 420
Db 361 TGCTGGGTAGCGCTACCCCGAGCTCGCAGCTAGGAAAGCCAGCGTCCCGCAGACA 420
Qy 421 ATACGACGCCACGTCGAT----- 438
Db 421 ATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG 480

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 19, 2003, 18:03:34 ; Search time 1558.97 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-25
Perfect score: 606
Sequence: 1 ATGTTGGTAGGTCATCGA.....TGCTCCGGATCCTCTAATAG 606

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues 45562784
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	68.2	11.3	488	9	AV755731
C 2	54.6	9.0	492	9	AV758366
C 3	41.6	6.9	502	12	BI879124
C 4	40.6	6.7	275	9	AV835132

5	40.6	6.7	402	9	AV392783	AV392783
6	40.6	6.7	551	9	AV392165	AV392165
7	40.6	6.7	552	12	BI996341	BI996341
8	40.6	6.7	584	12	BI727879	BI727879
9	40.4	6.7	1201	13	BI356664	BI356664
10	40.2	6.6	1162	12	BM918259	BM918259
C 11	40	6.6	1201	9	AL513886	AL513886
C 12	39	6.4	359	12	BU252669	BU252669
C 13	39	6.4	375	12	BJ246716	BJ246716
C 14	39	6.4	840	29	CC335916	CC335916
C 15	39	6.4	873	14	CD446071	CD446071
C 16	38.6	6.4	925	29	CNS0091P	AL053013
C 17	38.4	6.3	636	12	BI960110	BI960110
C 18	38.4	6.3	702	14	CD432549	BI960110
C 19	38.4	6.3	970	29	CNS010C9	BI960110
C 20	38.4	6.3	987	29	CNS015VX	AL098787
C 21	38.2	6.3	533	6	AU192776	AL105975
C 22	38.2	6.3	538	6	AU193705	AU192776
C 23	38.2	6.3	544	6	AU190971	AU193705
C 24	38.2	6.3	544	6	AU192419	AU190971
C 25	38.2	6.3	1270	12	BG968359	AU192419
C 26	38	6.3	354	14	CB966525	BG968359
C 27	38	6.3	1201	13	BI381961	CB966525
C 28	37.8	6.2	435	14	C72860	BI381961
C 29	37.8	6.2	533	29	CC010084	C72860
C 30	37.8	6.2	659	29	CC405164	CC010084
C 31	37.8	6.2	826	29	BZ736582	CC405164
C 32	37.8	6.2	895	29	CC359028	BZ736582
C 33	37.8	6.2	925	29	CC359026	CC359028
C 34	37.8	6.2	940	29	CC010085	CC359026
C 35	37.8	6.2	951	29	CC405167	CC010085
C 36	37.6	6.2	431	9	AV639153	CC405167
C 37	37.4	6.2	360	9	AJ473805	AV639153
C 38	37.4	6.2	637	13	BQ293470	AJ473805
C 39	37.4	6.2	641	13	BQ172543	BQ293470
C 40	37.4	6.2	650	14	CA828039	BQ172543
C 41	37.4	6.2	834	29	BZ641450	CA828039
C 42	37.4	6.2	841	29	BZ641457	BZ641450
C 43	37.4	6.2	856	29	BZ578381	BZ641457
C 44	37.4	6.2	872	29	BZ555011	BZ578381
C 45	37.2	6.1	563	10	BE490055	BZ555011

ALIGNMENTS

RESULT 1	AV755731/c	488 bp	mRNA	linear	EST 19-OCT-2000
LOCUS	AV755731	BM Homo sapiens	cDNA clone	BMFAKB03 5', mRNA	sequence.
DEFINITION	AV755731				
ACCESSION	AV755731				
VERSION	AV755731.1	GI:10913579			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 488)				
AUTHORS	Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,X., Li,N., Qian,B., Liu,P., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.				
TITLE	Homo sapiens cDNA BM clones				
JOURNAL	Unpublished				
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai. Location/Qualifiers				

FEATURES

[illegible]

Db	484	GGCCATGAGCATGTGAAGCCACCACTGATGGCATGGTTTCGCCAGACCCGAAAGTTGTTG	425
Qy	360	CTGCTGGGTAGCGCTCACCCCCACCGCTCGCAGCTAGGAAGCCAGCGCTCCCCACACGAC	419
Db	424	CCAGTTTTCAGTCCACCATCCAGCGCTCAATGCATGTACTCGAGCGACCATAGCCCAT	365
Qy	420	AATACGACGCCAGCTCGATTCCGAGCTGTTTCAACATCTCGCCTCGCGGATGAGACGGT	479
Db	364	CTCACCGCGAGGGTGGCTAAGTAGCCTTACAGTTTCAGGTTTCAGTCTAGTCGGGAGAACACGAA	305
Qy	480	GCAGGACTGCAA	491
Db	304	GTAGCCGTGCAA	293

RESULT 4	
LOCUS	AV835132
DEFINITION	AV835132 K. Sato unpublished cDNA library: Hordeum vulgare subsp. spontaneum top three leaves adult, heading stage Hordeum vulgare subsp. spontaneum cDNA clone bah24018, mRNA sequence.
ACCESSION	AV835132
VERSION	AV835132.1 GI:14527221
KEYWORDS	EST.
SOURCE	Hordeum vulgare subsp. spontaneum
ORGANISM	Hordeum vulgare subsp. spontaneum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum. 1 (bases 1 to 275)
REFERENCE	Sato, K.
AUTHORS	Barley EST sequencing project in NIG and Okayama Univ
TITLE	Unpublished
JOURNAL	Contact: Kazuhiro Sato;
COMMENT	Research Institute for Bioresources Okayama University, Barley Germplasm Center Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan Email: kzsato@rib.okayama-u.ac.jp, URL: http://www.rib.okayama-u.ac.jp/barley/ Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct submission; database: http://www.shigen.nig.ac.jp/barley/Barley.html. Location/Qualifiers 1. .275 /organism="Hordeum vulgare subsp. spontaneum" /mol_type="mRNA" /cultivar="H602" /db_xref="taxon:77009" /clone="bah24018" /tissue_type="top three leaves" /dev_stage="adult, heading stage" /clone_lib="K. Sato unpublished cDNA library: Hordeum vulgare subsp. spontaneum top three leaves adult, heading stage" 30 a 123 c 67 g 46 t 9 others
BASE COUNT	30 a 123 c 67 g 46 t 9 others
ORIGIN	

Query Match	6.7%;	Score 40.6;	DB 9;	Length 275;
Best Local Similarity	48.2%;	Pred. No. 4.3;		
Matches 109;	Conservative 0;	Mismatches 117;	Indels 0;	Gaps 0;

Qy	306	GATCATGCACACCCCGGGTGGTGGCCCTTCGCTTCGGGAGAACAACTTCCCGCTGCTG	365
Db	1	GGTCTCGACGGAAACNCCGCCCTTCCTCGCTTCGTTGCTCCCTGCGCTGGCACCGCCG	60
Qy	366	GGTAGCGCTCACCCCCACCGCTCGCAGCTAGGAAGCCAGCGGTCCCAACACGACAAATACG	425
Db	61	CGGCCAGGGCGCACTCACTTCTGTCTCGTTGCGCTTCCCTCTCGCGCCCCCTCTG	120
Qy	426	ACGCCAGTCGATTCCGAGCTGTTTCAACATCTCGCCTCGCGGATGAGACGGTGCAGGA	485
Db	121	ACGCTGGCGTACGCTGGGTGAGTTACATCTCTCGTCTCGGAGCGGACGCTGCGCGA	180

Mon Dec 22 13:28:43 2003

Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

RESULT 6
AV392165
LOCUS
DEFINITION
AV392165 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
CDNA clone CM083e05_x 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 551)
Asanizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags
DNA Res. 6 (6), 369-373 (1999)
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..551
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM083e05_r"
/dev_stage="photoautotrophic growth"
/clone_lib="Chlamydomonas reinhardtii C9"
/notes="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 94 a 182 c 189 g 85 t 1 others

Query Match 6.7%; Score 40.6; DB 9; Length 551;
Best Local Similarity 45.3%; Pred. No. 5.2;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
Qy 42 CGACCTCGTGGGGTACATTCGGCTCGCGCGCCCTAGGGGGCGCTGCCAGGGCCCT 101
Db 108 CGAGCTCATCTGTCATTTGTCGCGGCACTGCCAACATGAAGACGTGCTGACGGACCT 167
Qy 102 GCGGCATGGCGTCCGGGTTCTGGAGGACGGGTGAACATATGCACAGGGAATTTGCCCG 161
Db 168 GCGCGCGCGCGCGGAGTGGAGGGCGGTACGCGACGAGTCCGTGAGCTTGGCGC 227
Qy 162 TTGCTCTTCTATCTTCCTTGGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 221
Db 228 CCGCAAGGTGTTGACGAGATCAAGAGTAGCTGCTGTAACCTCAAGGCCCAAGAACCCAG 287
Qy 222 TTATGAAGTGGGCAACGTGTCCGGATGATACCATGTACAGAACGATGCTCCAACTCAAG 281
Db 288 CTTGCGCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 347
Qy 282 CATTGTGTATGAGGACGGGACATGATCATGCACACCCCGGGTCCGTGCGCTGCGTTG 341
Db 348 CCGTATGACCAACGACGAGGAGTTTGGCGGGCGCATCTACGGCGGCTGCGCCATGCCGG 407
Qy 342 CGAGAACAACTCTTCCCGCTGTGGGT 368
Db 408 CAAGAAGACCAAGGGCAGCTACATGAT 434

RESULT 7
BI996341
LOCUS
DEFINITION
1031037A07.y2 C. reinhardtii CC-1690, Stress II (normalized),
EST 25-OCT-2001

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 552)
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers
1..552
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
Polya mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into Lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with Exassist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
Research 6: 791-806." 86 t

BASE COUNT 93 a 184 c 189 g
ORIGIN
Query Match 6.7%; Score 40.6; DB 12; Length 552;
Best Local Similarity 45.3%; Pred. No. 5.2;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
Qy 42 CGACCTCGTGGGGTACATTCGGCTCGCGCGCCCTAGGGGGCGCTGCCAGGGCCCT 101
Db 110 CGAGCTCATCTGTCATTTGTCGCGGCACTGCCAACATGAAGACGTGCTGACGGACCT 169
Qy 102 GCGGCATGGCGTCCGGGTTCTGGAGGACGGGTGAACATATGCACAGGGAATTTGCCCG 161
Db 170 GCGCGCGCGCGCGGAGTGGAGGGCGGTACGCGACGAGTCCGTGAGCTTGGCGC 229
Qy 162 TTGCTCTTCTATCTTCCTTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 221
Db 230 CCGCAAGGTGTTGACGAGATCAAGAGTAGCTGCTGTAACCTCAAGGCCCAAGAACCCAG 289
Qy 222 TTATGAAGTGGGCAACCTGTCCGGATGATACCATGTACAGAACGATGCTCCAACTCAAG 281
Db 290 CTTGCGCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 349
Qy 282 CATTGTGTATGAGGACGGGACATGATCATGCACACCCCGGGTCCGTGCGCTGCGTTG 341
Db 350 CCGTATGACCAACGACGAGGAGTTTGGCGGGCGCATCTACGGCGGCTGCGCCATGCCGG 409


```

QY 342 GGAGAACAACTCTTCCCGTCTGGGT 368
Db 410 CAAGAAGACCAAGGCGAGCTACATGAT 436

RESULT 8
BI727879
LOCUS 584 bp mRNA linear EST 19-SEP-2001
DEFINITION 1031095C12.y1 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BI727879
VERSION BI727879.1 GI:15703574
KEYWORDS EST
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

FEATURES
source
1..584
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
BASE COUNT 106 a 188 c 197 g 93 t
ORIGIN
Query Match 6.7%; Score 40.6; DB 12; Length 584;
Best Local Similarity 45.3%; Pred. No. 5.3;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGGTACATTCGCTCGCGGCCCCCTAGGGGGCGTCCAGGCGCCT 101
Db 47 CGAGTCATCTTGTCATTCGCGGCGACTGCCACATGAAGAGCGTCTCAGCAGCT 106

QY 102 GCGCATGCGCTCCGGGTTCTCGAGACCGCGTGAACATAATGCAACAGGAAATTCGCCGG 161
Db 107 GCGCGCGCGCGCGAGTGGAGGCGCGCTACGCGCACGAGTCGTCGAGCTTGGGCG 166

QY 162 TTGCTCTTCTTAATCTCTTGGCTTTGCTGCTGCTGACCGTTCCAGCTTCGCGC 221
Db 167 CGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCCAAGACCCAG 226

QY 222 TTATGAGTGGCAACGTCGTCGGGATGTACCATGTACGAACGACTGCTCCAACCTCAAG 281
Db 227 CTTGCGCGTCCGTCGCTGGGCGCACTCGCTGGGCGGCGCACCGCGGCTGCTGTCGAT 286

QY 282 CATTGTGTATGAGCGAGCGGACATGATCATGCACACCCCGGGTGGCTGCTGCTGTCG 341
Db 287 CTTGATGACACGACGAGGAGTTTGGCGGCGCATCTACGCGGCGGTGCCCATGCCGGG 346

QY 342 GGAGAACAACTCTTCCCGTCTGGGT 368
Db 347 CAAGAAGACCAAGGCGAGCTACATGAT 373

RESULT 9
BI727879
LOCUS 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS01015YB03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664.1 GI:30378083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS01015CA02NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01015YB03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN
Query Match 6.7%; Score 40.4; DB 13; Length 1201;
Best Local Similarity 10.4%; Pred. No. 7.4;
Matches 52; Conservative 233; Mismatches 212; Indels 3; Gaps 1;

QY 34 GGCCTTCGCGACCTCGTGGGTACATTCGCTCGCGGCCCCCTAGGGGGCGCTGCC 93
Db 618 GNTNTSSSSSTNNNNSSSSNNNTNTTBTBTSSSSSTSSSTSBTSBTSBTSST 677

QY 94 AGGCGCCTCGGCGCATCGGCTCGGCTTCTGGAGGACGCGTGAACATGTCACACAGGAAT 153
Db 678 SSSSSSBTTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 737

QY 154 TTGCCCGGTGCTCTTCTTCTATCTCTTCTTGGCTTGTCTGCTGCTGACCGTTCCA 213
Db 738 TTKSSSSTBSSTSTTBTSTCTTTBTSTBTBTSSSYBSBSBTSSCSTSSBTTSS 797

QY 214 GCTTCGCTTATGAAGTGGCAACGTCGCGGGATGTACCATGTACGACACACTGCTCC 273
Db 798 TBSMTSSSBTCTSSSSSSSSBTTSTSTTTSTKSSBTT---SBBTBSSSSTSTTWTBTSC 854

```


[illegible]

RESULT 10	BM918259	linear	1162 bp	mRNA	EST 12-MAR-2002
LOCUS	BM918259	AGENCOURT 6611605 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485649			
DEFINITION		5', mRNA sequence.			
ACCESSION	BM918259	BM918259.1	GI:19368638		
VERSION		EST.			
KEYWORDS		Homo sapiens (human)			
SOURCE		Homo sapiens			
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE		NIH-MGC http://mgc.nci.nih.gov/ .			
AUTHORS		National Institutes of Health, Mammalian Gene Collection (MGC)			
TITLE		Unpublished			
JOURNAL		Contact: Robert Strausberg, Ph.D.			
COMMENT		Email: cgapbs-r@mail.nih.gov			
		Tissue Procurement: Dr. Daniel McVicar, DBS/NCI			
		cDNA Library Preparation: Rubin Laboratory			
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
		DNA Sequencing by: Agencourt Bioscience Corporation			
		Clone distribution: MGC clone distribution information can be			
		found through the I.M.A.G.E. Consortium/LLNL at:			
		http://image.llnl.gov			
		Plate: LLCM2016	row: n	column: 18	
		which quality sequence stop: 567.			

```

FEATURES
source
      Location/Qualifiers
1. ..1162
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:5485649"
   /tissue_type="natural killer cells, cell line"
   /lab_host="DH10B (phage-resistant)"
   /clone_lib="NIH MGC 106"
   /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
   EcoRI; cDNA made by oligo-dT priming. Directionally cloned
   into EcoRI/XhoI sites using the following 5' adaptor:
   GGACACGAG(G). Library constructed by Ling Hong in the
   laboratory of Gerald M. Rubin (University of California,
   Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
   Superscript II RT (Life Technologies). Note: this is a
   NIH MGC Library."
BASE COUNT      224 a  _499 c      240 g      198 t      1 others
ORIGIN
      Query Match      6.6%; Score 40.2; DB 12; Length 1162;
      Best Local Similarity 54.4%; Pred No. 8,2;
      Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
      313 CACACCCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTCCCGCTGCTGGTAGCG 372
Qy

```

Db	715	CCCCCCCCCGGGTCCCTGTGCCACACCCCGGAGCCCCAAACCCCCCGGGACCGCTCC	774
Qy	373	CTCACCCCGACGCTCGGAGCTAGGAAGCCAGCGTCCCAACACGACGAATAACGACGCCAC	432
Db	775	CCCACTGCACCGCAGCACCCGCCCATATCGCCGCCCTACCCGATCACCCTACGCCAC	834
Qy	433	GTTCGATTCGCCAGCTGTTCACCATCGCC	461
Db	835	GCCTGATCCCGGCCCTGCACACCCCGCC	863

RESULT 11	AL513886	1201 bp	linear	EST 08-MAY-2003
LOCUS	AL513886/c			
DEFINITION	AL513886 Homo sapiens PLACENTA Homo sapiens cDNA clone			Cl08A006ZG08
	5-PRIME, mRNA sequence.			
ACCESSION	AL513886			
VERSION	AL513886.2	GI:30463771		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1201)			
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
COMMENT	Unpublished			
	On Feb 13, 2001 this sequence version replaced gi:12777380.			
	Contact: Genoscope			
	Genoscope - Centre National de Sequencage			
	BP 191 91006 EVRY cedex - France			
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
	Library was constructed by Life Technologies, a division of			
	Invitrogen. This sequence belongs to sequence cluster 4924.f For			
	more information about this cluster, see			
	http://www.genoscope.cns.fr/			
	cgi-bin/cluster.cgi?seq=Cl08A006ZG08RPl&cluster=4924.f. Contact :			
	Feng Liang Email : fliang@lifetech.com URL :			
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
	Paraday Avenue Genoscope sequence ID : Cl08A006ZG08RPl.			
	Location/Qualifiers			

FEATURES
source

```

1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA006ZG08"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched, into
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
201 a 311 c 349 g 146 t 194 others
BASE COUNT
ORIGIN

Query Match 6.6%; Score 40; DB 9; Length 1201;
Best Local Similarity 26.5%; Pred. No. 9.4;
Matches 103; Conservative 104; Mismatches 179; Indels 3; Gaps 1;

QY 32 GCGGCTTCGCGACCTCGTGGGTACATTCGCTGTCGGCGCCCCCTAGGGGGCGCTG 91
Db 1129 GSGCGTGGGCSBSCNTYKKGKBSSSSSGSGSGSSCGCGGGGGCGGGG 1070
QY 92 CCAGGGCCCTGGCGATGCGCTCCGGGTTCTGGAGGACGGCGTAACTATGCAACAGGA 151
Db 1069 CCSGCKGGGGKSSSGSGGCCCGCGG---GGGSSGSSGGGSCCWAASSSAYKKGK 1013
QY 152 ATTGCCCCGTGCTCTTCTATCTTCTCTCTTGGCTTTGCTGCTGTGTGACCGTTC 211
Db 1012 GKKTCTTTTMMAAATMTHTTTTWTWTYTTCTAAGGGGTAKCVAKCCCMCCCCAMGCTS 953

```



```

QY 212 CAGCTTCGGCTTATGAGTGGCAACGCTGTCGGGATGTACCATGTGCACGACGACTGCT 271
Db 952 GACSCSCCGCAADAACGAGMDSGAMKGTGVSCTTSRCRWGGTTSMMVGCATTYMM 893

QY 272 CCAACTCAAGCATTTGTATGAGCGAGCATATGATCATGCACACCCCGCGTGGTCC 331
Db 892 AVBSYTGTRWTWTGTSBTCTYASGSGMYSSKKBKCKCMAYAACSCAGASCST 833

QY 332 CTGCGCTTCGGAGAACAACTCTTCCTCGCTGCTGGTAGCGTCACTCCACCGCTGCAG 391
Db 832 SGCSRGKTKTKTTGCTGCTGTAAGASMBAGTWAGGGGGGGCCCGCCYCSMCCCGCCCVB 773

QY 392 CTAGGAACGCCAGCGTCCCCACACGACA 420
Db 772 BBCCMCWCHTKCKMCKCRGACTYCCCA 744

RESULT 12
BU252669/c
LOCUS
DEFINITION
Triticum aestivum (bread wheat)
ACCESSION
BU252669
VERSION
BU252669.1 GI:20061830
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 359)
Ogihara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stages="Feekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
BASE COUNT 70 a 115.c 107 g 67 t
ORIGIN
1..359
Query Match 6.4%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTCGGGACCTCGTGGGTACATTCCGCTCGTGGCGCCCGCTAGG 83
Db 297 CTTCAAGTGCACACGCGCTCTGGAAGCGCTCAGGGCGGTTCGACGCGCTCGCGTCGG 238

QY 84 GGGCGCTGCAGGCGCCCTGGCGCATGGCTCGGGTTCTGGAGGACGGCGTCAACTATG 142
Db 237 GGACGCCGACGCCCTGGCGCAGGACGTCACGTCGTGCGGTGCACGTGCCCAAGG 179

RESULT 13
BU246716
LOCUS
DEFINITION
Triticum aestivum (bread wheat)
ACCESSION
BU246716
VERSION
BU246716.1 GI:20058228
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 375)
Ogihara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..375
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stages="Feekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
BASE COUNT 81 a 107 c 110 g 77 t
ORIGIN
1..375
Query Match 6.4%; Score 39; DB 12; Length 375;
Best Local Similarity 58.0%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTCGGGACCTCGTGGGTACATTCCGCTCGTGGCGCCCGCTAGG 83
Db 36 CTTCAAGTGCACACGCGCTCTGGAAGCGCTCAGGGCGGTTCGACGCGCTCGCGTCGG 95

QY 84 GGGCGCTGCAGGCGCCCTGGCGCATGGCTCGGGTTCTGGAGGACGGCGTCAACTATG 142
Db 96 GGACGCCGACGCCCTGGCGCAGGACGTCACGTCGTGCGGTGCACGTGCCCAAGG 154

RESULT 14
CC335916
LOCUS
DEFINITION
OGUJ60TV ZM 0.7.1.5_KB Zea mays genomic clone ZMMBma0393124,
genomic survey sequence.
ACCESSION
CC335916
VERSION
CC335916.1 GI:30805329
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek
, R.W., Nurnberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..840

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 43.7045 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-25
Perfect score: 606
Sequence: 1 ATGTTGGTGAAGTCATCGA.....TGCTCGGATCCTCTAATAG 606

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	606	100.0	606	3	US-08-612-973-25
2	606	100.0	606	3	US-08-927-597-25
3	598.2	98.7	723	3	US-08-612-973-21
4	598.2	98.7	723	3	US-08-927-597-21
5	597.4	98.6	636	3	US-08-612-973-27
6	597.4	98.6	636	3	US-08-927-597-27
7	556	91.7	561	3	US-08-612-973-23
8	556	91.7	561	3	US-08-927-597-23
9	516.2	85.2	795	3	US-08-612-973-5
10	516.2	85.2	795	3	US-08-927-597-5
11	513.2	84.7	2082	3	US-08-612-973-47
12	513.2	84.7	2082	3	US-08-927-597-47
13	513.2	84.7	2433	3	US-08-612-973-49
14	513.2	84.7	2433	3	US-08-927-597-49
15	456.4	75.3	633	3	US-08-612-973-7
16	456.4	75.3	633	3	US-08-927-597-7
17	451.6	74.5	636	3	US-08-612-973-13
18	451.6	74.5	636	3	US-08-927-597-13
19	449.2	74.1	1539	2	US-08-470-426B-17
20	449.2	74.1	1539	2	US-08-470-426B-14
21	444.4	73.3	9595	3	US-09-014-416-4
22	444.4	73.3	9599	3	US-09-014-416-6
23	442.8	73.1	932	1	US-08-081-072-15
24	442.8	73.1	932	1	US-08-449-093A-15
25	442.8	73.1	2116	3	US-08-191-160-21
26	441.2	72.8	9472	4	US-08-150-204B-96
27	441	72.8	1037	1	US-08-462-195-1

28	441	72.8	1037	2	US-08-636-883-1	Sequence 1, Appli
29	441	72.8	1037	3	US-09-127-829-1	Sequence 1, Appli
30	430	71.0	1167	1	US-08-324-977-9	Sequence 9, Appli
31	430	71.0	1167	2	US-08-384-616-9	Sequence 9, Appli
32	430	71.0	1167	2	US-08-904-686A-9	Sequence 9, Appli
33	430	71.0	1167	3	US-09-315-850-9	Sequence 9, Appli
34	430	71.0	1499	1	US-08-324-977-3	Sequence 3, Appli
35	430	71.0	1499	2	US-08-384-616-3	Sequence 3, Appli
36	430	71.0	1499	2	US-08-904-686A-3	Sequence 3, Appli
37	430	71.0	1499	3	US-09-315-850-3	Sequence 3, Appli
38	430	71.0	6039	1	US-08-324-977-11	Sequence 11, Appli
39	430	71.0	6039	2	US-08-384-616-11	Sequence 11, Appli
40	430	71.0	6039	2	US-08-904-686A-11	Sequence 11, Appli
41	430	71.0	9030	3	US-09-315-850-11	Sequence 11, Appli
42	430	71.0	9030	1	US-08-324-977-13	Sequence 13, Appli
43	430	71.0	9030	2	US-08-384-616-13	Sequence 13, Appli
44	430	71.0	9030	2	US-08-904-686A-13	Sequence 13, Appli
45	430	71.0	9030	3	US-09-315-850-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-08-612-973-25
; Sequence 25, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..603
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..600

	541	541	541	601	601	600
Qy	ATGATGATGAACTGGTGCCTTACAA	CGGCGCTGGTGGTATCGACAGTCTCGGATCCTC	600			
Db	ATGATGATGAACTGGTGCCTTACAA	CGGCGCTGGTGGTATCGACAGTCTCGGATCCTC	600			
Qy	TAATAG	606				
Db	TAATAG	606				

RESULT 3
 US-08-612-973-21
 ; Sequence 21, Application US/08612973
 ; Patent No. 6150134
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: MAERTENS, GEERT
 ; APPLICANT: BOSMAN, FONS
 ; APPLICANT: DE MARTYNOFF, GUY
 ; APPLICANT: BUYSE, MARIE-ANGE
 ; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
 ; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
 ; NUMBER OF SEQUENCES: 111
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: NIXON & VANDERHYE P. C.

	Query Match	98.7%	Score 598.2	DB 3	Length 723
	Best Local Similarity	99.5%	Pred. No. 1.1e-150		
	Matches 600	Conservative 0	Mismatches 3	Indels 0	Gaps 0
Qy	1	ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGCTGGGGGTACATT	60		
Db	1	ATGTTGGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGCTGGGGGTACATT	60		
Qy	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCTCCGGGTT	120		
Db	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCTCCGGGTT	120		

QY	121	CTGAGAGACGGCGTGAAC	TATGCAACAGGGAATTTG	CCCGGTTGCTCTTTCT	TATCTTC	180
Db	121	CTGAGAGACGGCGTGAAC	TATGCAACAGGGAATTTG	CCCGGTTGCTCTTTCT	TATCTTC	180
QY	181	CTCTTGCTTTTGCTGCTG	CTGACAGTTTCCAGCTTT	CCGCTTATGAAGTGGGA	ACCTG	240
Db	181	CTCTTGCTTTTGCTGCTG	CTGACAGTTTCCAGCTTT	CCGCTTATGAAGTGGGA	ACCTG	240
QY	241	TCCGGGATGTACCATGT	CACGAACGACTGTCTCAA	CTCAAGCATTTGTGTATG	GAGCAGCG	300
Db	241	TCCGGGATGTACCATGT	CACGAACGACTGTCTCAA	CTCAAGCATTTGTGTATG	GAGCAGCG	300
QY	301	GACATGATCATGACAC	CCCCGGGTGGTGCCCTG	CGTTCGGGAGAACAACT	CTTCCCGC	360
Db	301	GACATGATCATGACAC	CCCCGGGTGGTGCCCTG	CGTTCGGGAGAACAACT	CTTCCCGC	360
QY	361	TGCTGGGTAGCGCTCAC	CCCCAGCTCGCAGCTAG	GAAAGCCAGCTCCCA	CACGACA	420
Db	361	TGCTGGGTAGCGCTCAC	CCCCAGCTCGCAGCTAG	GAAAGCCAGCTCCCA	CACGACA	420
QY	421	ATACGACGCCACGCTCG	ATTTCCAGCTGTTTCA	CCACTCTCGCTCCGG	GCATGAGACGGTG	480
Db	421	ATACGACGCCACGCTCG	ATTTCCAGCTGTTTCA	CCACTCTCGCTCCGG	GCATGAGACGGTG	480
QY	481	CAGGACTGCAATTTGCT	CAATCTATCCGGGCCA	CAATAACGGGTACCG	GTATGGGTGGAT	540
Db	481	CAGGACTGCAATTTGCT	CAATCTATCCGGGCCA	CAATAACGGGTACCG	GTATGGGTGGAT	540
QY	541	ATGATGATGAACGTGCT	CGCCTACAAAGCCCTG	TGCTGGTATCGCAGCT	GCTCCGGATCCCA	600
Db	541	ATGATGATGAACGTGCT	CGCCTACAAAGCCCTG	TGCTGGTATCGCAGCT	GCTCCGGATCCCA	600
QY	601	TAA	603			
Db	601	CAA	603			

RESULT 4
 US-08-927-597-21
 ; Sequence 21, Application US/08927597
 ; Patent No. 6245503
 ; GENERAL INFORMATION:
 ; APPLICANT: MAERTENS, GEERT
 ; APPLICANT: BOSMAN, FONS
 ; APPLICANT: DE MARTYNOFF, GUY
 ; APPLICANT: BUYSSE, MARIE-ANGE
 ; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
 ; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
 ; NUMBER OF SEQUENCES: 111
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHYE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/927,597
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/612,973
 ; FILING DATE: 11-MAR-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BYRNE, THOMAS E.
 ; REGISTRATION NUMBER: 32,205
 ; REFERENCE/DOCKET NUMBER: 1487-10
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..720
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..717
US-08-927-597-21

CURRENT APPLICATION DATE: 11-MAR-1996
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..633
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..630
US-08-612-973-27

Query Match	98.6%;	Score 597.4;	DB 3;	Length 636;
Best Local Similarity	99.8%;	Pred. No. 1.7e-150;		
Matches 598;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
QY	1	ATGTTGGGTAAAGGTCAATCGATACCCCTTACATGCGGCTTCGCCACCTCGTGGGGGTACATT	60	
DB	1	ATGTTGGGTAAAGGTCAATCGATACCCCTTACATGCGGCTTCGCCACCTCGTGGGGGTACATT	60	
QY	61	CGGCTCGTGGCGGCCCCCTTAGGGGGCGCTGCCAGGCGCCCTGGCGCATGCGCTCCGGGTT	120	
DB	61	CGGCTCGTGGCGGCCCCCTTAGGGGGCGCTGCCAGGCGCCCTGGCGCATGCGCTCCGGGTT	120	
QY	121	CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTC	180	
DB	121	CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTC	180	
QY	181	CTCTTTGGCTTTGCTGCTCTGTGTCAGCGTTCCAGCTTCCGCTTATGAAGTGCACAAAGTG	240	
DB	181	CTCTTTGGCTTTGCTGCTCTGTGTCAGCGTTCCAGCTTCCGCTTATGAAGTGCACAAAGTG	240	
QY	241	TCCGGGATGTACATGTACAGAAACGACTGTCTCCAACTCAAGCATTTGTATGAGCGACGC	300	
DB	241	TCCGGGATGTACATGTACAGAAACGACTGTCTCCAACTCAAGCATTTGTATGAGCGACGC	300	
QY	301	GACATGATCATGCACACCCCGGGTGGTGCCCTCGGTTCCGGAGAACAACTCTTCCCGC	360	

Db 301 GACATGATCATGACACCCCGGGTGGTGGCTGGCTGGGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGCTACCCCGGAGCTGCGAGCTAGGAGAGCCAGCTGCCACACGACA 420
Db 361 TGTGGGTAGCGCTACCCCGGAGCTGCGAGCTAGGAGAGCCAGCTGCCACACGACA 420
QY 421 ATAGGAGCGGAGCTGATTCGAGCTGTTTACCATCTCGCTCGCGGATGAGAGCGTG 480
Db 421 ATAGGAGCGGAGCTGATTCGAGCTGTTTACCATCTCGCTCGCGGATGAGAGCGTG 480
QY 481 CAGGAGCTGCAATTTGCTCAATCTATCCCGGACACATAACCGGTCACCGTATGGCTTGGGAT 540
Db 481 CAGGAGCTGCAATTTGCTCAATCTATCCCGGACACATAACCGGTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAAGCTGCTGCGCTACACGCGGCTGGTGGTATCGAGCTGCTCGGATCCT 599
Db 541 ATGATGATGAAGCTGCTGCGCTACACGCGGCTGGTGGTATCGAGCTGCTCGGATCCT 599

RESULT 6

US-08-927-597-27
; Sequence 27, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630

US-08-927-597-27

Query Match 98.6%; Score 597.4; DB 3; Length 636;
Best Local Similarity 99.8%; Pred. No. 1.7e-150;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAAGTCAATCCCTTACATCGGGCTTCGCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCAATCCCTTACATCGGGCTTCGCGACCTCGTGGGGTACATT 60
QY 61 CGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGAGCCCTGGCGCATGGCGTCCGGTT 120
Db 61 CGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGAGCCCTGGCGCATGGCGTCCGGTT 120
QY 121 CTGGAGGAGCGGTGAACCTATGCAACAGGGAATTTCCCGGTTGCTCTTTCTATCTTC 180
Db 121 CTGGAGGAGCGGTGAACCTATGCAACAGGGAATTTCCCGGTTGCTCTTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGACCGGTTTCAGGCTTCGGCTTATCAAGTGGCGCAACGTG 240
Db 181 CTCTTGGCTTTGCTGCTGCTGACCGGTTTCAGGCTTCGGCTTATCAAGTGGCGCAACGTG 240
QY 241 TCGGGATGTACCATGTCTACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGC 300
Db 241 TCGGGATGTACCATGTCTACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGC 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGCTACCCCGGAGCTCGAGCTAGGAAGCCAGGCTCCCGACGACA 420
Db 361 TGTGGGTAGCGCTACCCCGGAGCTCGAGCTAGGAAGCCAGGCTCCCGACGACA 420
QY 421 ATACGAGCGGACGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 480
Db 421 ATACGAGCGGACGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 480
QY 481 CAGGAGCTGCAATTTGCTCAATCTATCCCGGACACATAACCGGTCACCGTATGGCTTGGGAT 540
Db 481 CAGGAGCTGCAATTTGCTCAATCTATCCCGGACACATAACCGGTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAAGCTGCTGCGCTACACGCGGCTGGTGGTATCGAGCTGCTCCGGATCCT 599
Db 541 ATGATGATGAAGCTGCTGCGCTACACGCGGCTGGTGGTATCGAGCTGCTCCGGATCCT 599

RESULT 7

US-08-612-973-23
; Sequence 23, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996

RESULT 8
US-08-927-597-23
; Sequence 23, Application US/08927597

Mon Dec 22 13:28:42 2003

COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: US 08/612,973
APPLICATION NUMBER:
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789
US-08-927-597-5

Query Match 85.2%; Score 516.2; DB 3; Length 795;
Best Local Similarity 88.9%; Pred. No. 8.5e-129;
Matches 600; Conservative 0; Mismatches 3; Indels 72; Gaps 1;
QY 1 ATGTGGGTAAGGTGATCATACCTTATCATCGGGCTTCGCGGACCTCGTGGGTACATT 60
Db 1 ATGTGGGTAAGGTGATCATACCTTATCATCGGGCTTCGCGGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGGGCATGGCGTCCGGTT 120
Db 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGGGCATGGCGTCCGGTT 120
QY 121 CTGGAGGAGCGGTGAACATATCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTC 180
Db 121 CTGGAGGAGCGGTGAACATATCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTC 180
QY 181 CTCTGGCTTTGTCTGTCTGTACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
Db 181 CTCTGGCTTTGTCTGTCTGTACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
QY 241 TCCGGGATGATCATGTACAGCAAGCTGCTCAACTCAAGCAATTTGTATGAGCAGCG 300
Db 241 TCCGGGATGATCATGTACAGCAAGCTGCTCAACTCAAGCAATTTGTATGAGCAGCG 300
QY 301 GACATGATCATGACACCCCGGGTGGTCCCTGCTTCCGGAGAACACTCTTCCCGC 360
Db 301 GACATGATCATGACACCCCGGGTGGTCCCTGCTTCCGGAGAACACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCAAGCTAGGAAGCCAGCGTCCCGACCGACA 420
Db 361 TGCTGGGTAGCGCTCACCCCAAGCTAGGAAGCCAGCGTCCCGACCGACA 420
QY 421 ATACGACGCCACGTCGAT----- 438
Db 421 ATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG 480

QY 439 -----TCCGAGCTGTTCCACCATCTCGCTCGCGG 468
Db 481 GGGGACCTCTGGGATCTGTCTTCTCTCCGAGCTGTTCCACCATCTCGCTCGCGG 540
QY 469 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT 528
Db 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT 600
QY 529 ATGGCTTGGGATATGATGATGAACCTGGTGCCTTACACGGCCCTGGTGTATCGCAGCTG 588
Db 601 ATGGCTTGGGATATGATGATGAACCTGGTGCCTTACACGGCCCTGGTGTATCGCAGCTG 660
QY 589 CTCGGATCTCTAA 603
Db 661 CTCGGATCCCAAA 675

RESULT 11
US-08-612-973-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-612-973-47

Query Match 84.7%; Score 513.2; DB 3; Length 2082;
Best Local Similarity 88.8%; Pred. No. 7.1e-128;
Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;

QY 4 TTGGGTAAGTGCATGATACCTTACATGCGGCTTCGCCGACTCGTGGGTACATTCGG 63
DB 4 TTGGGTAAGTGCATGATACCTTACATGCGGCTTCGCCGACTCGTGGGTACATTCGG 63
QY 64 CTCGTCGGCGCCCTAGGCGGCGCTGCAGGCGCTCGCGCATGCGGCTCGGGTTCTG 123
DB 64 CTCGTCGGCGCCCTAGGCGGCGCTGCAGGCGCTCGCGCATGCGGCTCGGGTTCTG 123
QY 124 GAGGACGGCGTGAACATATCAACAGGGAATTTGCCGGTTGCTTCTTCTATCTTCCTC 183
DB 124 GAGGACGGCGTGAACATATCAACAGGGAATTTGCCGGTTGCTTCTTCTATCTTCCTC 183
QY 184 TTGGCTTTGCTGCTGCTGATGACGCTTCAGCTTCGCTTATGAAGTGCAGGCTTCG 243
DB 184 TTGGCTTTGCTGCTGCTGATGACGCTTCAGCTTCGCTTATGAAGTGCAGGCTTCG 243
QY 244 GGGATGTACCATGTACGAACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
DB 244 GGGATGTACCATGTACGAACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
QY 304 ATGATCATGACACACCCCGGCTGCGTTCGCTTCGGGAGAACAACTCTTCCCGCTGC 363
DB 304 ATGATCATGACACACCCCGGCTGCGTTCGCTTCGGGAGAACAACTCTTCCCGCTGC 363
QY 364 TGGGTAGCGCTACCCCGACGCTCGAGCTAGGAGCGGCTCCCAACGACAATA 423
DB 364 TGGGTAGCGCTACCCCGACGCTCGAGCTAGGAGCGGCTCCCAACGACAATA 423
QY 424 CGACGCCACGTCGAT----- 438
DB 424 CGACGCCACGTCGATTTGCTGCGGCGGCTGCTTCTGTCGCTATGATGCGGG 483
QY 439 -----TCCAGCTGTTCAACATCTCGCTCGCGGCAAT 471
DB 484 GACCTCTCGGATCTGCTTCTCGTCTCCAGCTGTTCAACATCTCGCTCGCGGCAAT 543
QY 472 GAGAGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGACACATAACCGGTATG 531
DB 544 GAGAGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGACACATAACCGGTATG 603
QY 532 GCTTGGGATATGATGAACTGCTGCTCAACGCGGCTCGTGTATGCGAGCTGCTC 591
DB 604 GCTTGGGATATGATGAACTGCTGCTCAACGCGGCTCGTGTATGCGAGCTGCTC 663
QY 592 CGGATCCTCTAA 603
DB 664 CGGATCCCAAA 675

RESULT 12
US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2079
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2076
US-08-927-597-47

Query Match 84.7%; Score 513.2; DB 3; Length 2082;
Best Local Similarity 88.8%; Pred. No. 7,1e-128;
Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;
QY 4 TTGGGTAAGTGCATGATACCTTACATGCGGCTTCGCCGACTCGTGGGTACATTCGG 63
DB 4 TTGGGTAAGTGCATGATACCTTACATGCGGCTTCGCCGACTCGTGGGTACATTCGG 63
QY 64 CTCGTCGGCGCCCTAGGCGGCGCTGCAGGCGCTCGCGCATGCGGCTCGGGTTCTG 123
DB 64 CTCGTCGGCGCCCTAGGCGGCGCTGCAGGCGCTCGCGCATGCGGCTCGGGTTCTG 123
QY 124 GAGGACGGCGTGAACATATCAACAGGGAATTTGCCGGTTGCTTCTTCTATCTTCCTC 183
DB 124 GAGGACGGCGTGAACATATCAACAGGGAATTTGCCGGTTGCTTCTTCTATCTTCCTC 183
QY 184 TTGGCTTTGCTGCTGCTGATGACGCTTCAGCTTCGCTTATGAAGTGCAGGCTTCG 243
DB 184 TTGGCTTTGCTGCTGCTGATGACGCTTCAGCTTCGCTTATGAAGTGCAGGCTTCG 243
QY 244 GGGATGTACCATGTACGAACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
DB 244 GGGATGTACCATGTACGAACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
QY 304 ATGATCATGACACACCCCGGCTGCGTTCGCTTCGGGAGAACAACTCTTCCCGCTGC 363
DB 304 ATGATCATGACACACCCCGGCTGCGTTCGCTTCGGGAGAACAACTCTTCCCGCTGC 363
QY 364 TGGGTAGCGCTACCCCGACGCTCGAGCTAGGAGCGGCTCCCAACGACAATA 423
DB 364 TGGGTAGCGCTACCCCGACGCTCGAGCTAGGAGCGGCTCCCAACGACAATA 423
QY 424 CGACGCCACGTCGAT----- 438
DB 424 CGACGCCACGTCGATTTGCTGCGGCGGCTGCTTCTGTCGCTATGATGCGGG 483
QY 439 -----TCCAGCTGTTCAACATCTCGCTCGCGGCAAT 471
DB 484 GACCTCTCGGATCTGCTTCTCGTCTCCAGCTGTTCAACATCTCGCTCGCGGCAAT 543
QY 472 GAGAGCGTGCAGGACTGCAATTTGCTCAATCTATCCCGGACACATAACCGGTATG 531


```

Db      544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTACCGGTATG 603
QY      532 GCTTGGGATATGATGATGAATGCTCGCCTCAACAGCGCCCTGGTGGTATCGCAGCTGCTC 591
Db      604 GCTTGGGATATGATGATGAATGCTGGTGGCTTCAACAGCGCCCTGGTGGTATCGCAGCTGCTC 663
QY      592 CGGATCCTCTTAA 603
Db      664 CGGATCCCAAA 675

RESULT 13
US-08-612-973-49
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2430
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2427
; US-08-612-973-49

Query Match 84.7%; Score 513.2; DB 3; Length 2433;
Best local Similarity 88.8%; Pred. No. 7.5e-128;
Matches 597; Conservative 3; Indels 72; Gaps 1;

QY      4 TTGGGTAAAGTCATGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63
Db      355 TTGGGTAAAGTCATGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 414
QY      64 CTGTCGGCGGCCCCCTAGGGGGCGCTGCGCAGGGCCCTGGCGCATGGCGTCCGGGTCTG 123

Db      415 CTCGTGCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTG 474
QY      124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCCCGGTTCCTTTCTCTATCTTCTCTC 183
Db      475 GAGGACGGCGTGAATATGCAACAGGGAATTTGCCCGGTTCCTTTCTCTATCTTCTCTC 534
QY      184 TTGGCTTTGCTGCTCTGCTGCTGACCGCTTCAGCTTCGGCTTATGAAGTCGCAACAGTGTCC 243
Db      535 TTGGCTTTGCTGCTCTGCTGCTGACCGCTTCAGCTTCGGCTTATGAAGTCGCAACAGTGTCC 594
QY      244 GGGATGTACCATGTTCACGAACGACTGCTCCAACTCAAGCAATTTGTATGAGGCAGCGGAC 303
Db      595 GGGATGTACCATGTTCACGAACGACTGCTCCAACTCAAGCAATTTGTATGAGGCAGCGGAC 654
QY      304 ATGATCATGACACACCCCGGGTTCGCTGCGCTTCGGGAGAAACAACCTCTTCCCGCTGC 363
Db      655 ATGATCATGACACACCCCGGGTTCGCTGCGCTTCGGGAGAAACAACCTCTTCCCGCTGC 714
QY      364 TGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAAGCCAGCGTCCCAACCAAGCAATA 423
Db      715 TGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAAGCCAGCGTCCCAACCAAGCAATA 774
QY      424 CGACGCCACGTCCAT 438
Db      775 CGACGCCACGTCCATTTGCTCGTTGGGGCGGTGCTTTCTGTTCCGCTATGTACGTGGGG 834
QY      439 -----TCCAGCTGTTTCCACCATCTCGCTCGCTCGCCGGCAT 471
Db      835 GACCTCTGGCGATCTGCTTCTTCCTCGCTCCAGCTGTTCCACCATCTCGCTCGCCGGCAT 894
QY      472 GAGACGGTGCAGGACTGCAATTTGCTCAATTTATCCCGGCCACATAAGCGGTACCGGTATG 531
Db      895 GAGACGGTGCAGGACTGCAATTTGCTCAATTTATCCCGGCCACATAAGCGGTACCGGTATG 954
QY      532 GCTTGGGATATGATGATGAATGCTGCTCGCTTACAAAGCCCTGCTGCTGATCGCAGCTGCTC 591
Db      955 GCTTGGGATATGATGATGAATGCTGCTCGCTTACAAAGCCCTGCTGCTGATCGCAGCTGCTC 1014
QY      592 CGGATCCTCTTAA 603
Db      1015 CGGATCCCAAA 1026

RESULT 14
US-08-927-597-49
; Sequence 49, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:
; LENGTH:
; TYPE:
; STRANDEDNESS:
; TOPOLOGY:
; MOLECULE TYPE:
; HYPOTHEtical:
; ANTI-SENSE:
; FEATURE:
; NAME/KEY:
; LOCATION:
; FEATURE:
; NAME/KEY:
; LOCATION:
; US-08-927-597-49
```


APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2430
NAME/KEY: mat_peptide
LOCATION: 1..2427
US-08-927-597-49

Query Match 84.7%; Score 513.2; DB 3; Length 2433;
Best Local Similarity 88.8%; Pred. No. 7.5e-128;
Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;
QY 4 TTGGGTAAGGTATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 63
DB 355 TTGGGTAAGGTATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 414
QY 64 CTGCTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTGCGCATGCGGTCCGGGTTCTG 123
DB 415 CTGCTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTGCGCATGCGGTCCGGGTTCTG 474
QY 124 GAGGACGGGTGAATCATGATCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTCCTC 183
DB 475 GAGGACGGGTGAATCATGATCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTGCTGACCGTTTCAGCTTCGCTTATGAAGTGGCAACGTTGTC 243
DB 535 TTGGCTTTGCTGCTGCTGACCGTTTCAGCTTCGCTTATGAAGTGGCAACGTTGTC 594
QY 244 GGGATGACCATGCTACGACGACTGCTCCAAGTCAAGCATTTGTTATGAGGCGGCGAC 303
DB 595 GGGATGACCATGCTACGACGACTGCTCCAAGTCAAGCATTTGTTATGAGGCGGCGAC 654
QY 304 ATGATCATGCACACCCCGGGTGGTGGCTTCCGAGAACAACTCTTCCCGCTGC 363
DB 655 ATGATCATGCACACCCCGGGTGGTGGCTTCCGAGAACAACTCTTCCCGCTGC 714
QY 364 TGGGTAGCGTCAACCCGACGCTCGAGCTAGGAAAGCCAGCGTCCCAACGACGACAA 423
DB 715 TGGGTAGCGTCAACCCGACGCTCGAGCTAGGAAAGCCAGCGTCCCAACGACGACAA 774
QY 424 CGACGCGCAGTCCGAT----- 438
DB 775 CGACGCGCAGTCCGATTTGCTGTTGGGGGGGCTGCTTCTTCTGCTGATGACGTTGGG 834
QY 439 -----TCCAGCTGTTACCATCTCGCTCGCTCGCGGCAT 471
DB 835 GACCTCTGGGATCTGCTTCTTCTCCAGCTGTTTACCATCTCGCTCGCGGCAT 894
QY 472 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTCAACGATG 531
DB 895 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTCAACGATG 954
QY 532 GCTTGGGATATGATGATGAACTGCTGCTTACAAAGCGGCTTGTGTTATGCGAGTCTC 591
DB 955 GCTTGGGATATGATGATGAACTGCTGCTTACAAAGCGGCTTGTGTTATGCGAGTCTC 1014

QY 592 CGGATCCTCTAA 603
DB 1015 CGGATCCCAAA 1026
RESULT 15
US-08-612-973-7
Sequence 7, Application US/08612973
Patent No. 6150134
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 633 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..630
NAME/KEY: mat_peptide
LOCATION: 1..627
US-08-612-973-7
Query Match 75.3%; Score 456.4; DB 3; Length 633;
Best Local Similarity 86.8%; Pred. No. 7.4e-113;
Matches 545; Conservative 0; Mismatches 11; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAAGGTATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAGGTATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CGCTCGTGGGCGCCCCCTAGGGGCGCTGCGAGGCGCTTGGCGCATGCGGTCCGGGTT 120
DB 61 CGCTCGTGGGCGCCCCCTAGGGGCGCTGCGAGGCGCTTGGCGCATGCGGTCCGGGTT 120
QY 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Mon Dec 22 13:28:42 2003

Qy	181	CTCTTGGCTTTGCTGTCCTGCTGACCGTTCCAGTTCGGCTTATGAAGTGGCAAGTG	240
Db	181		
Qy	241	TCGGGATGTACCATGTACAGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGCAGCG	300
Db	241		
Qy	301	GACATGATCATGCACACCCCGGGTGCCTGCTGGGAGAACAACTCTTCCCGC	360
Db	301		
Qy	361	TGCTGGGTAGCGCTACCCCGCAGCTCGCAGCTAGGAACGCCAGCGTCCCAACACGACA	420
Db	361		
Qy	421	ATACGACGCCACGTCGAT-----	438
Db	421		
Qy	439	-----TCCAGCTGTTACCATCTCGCTCGCGG	468
Db	439		
Qy	469	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGT	528
Db	469		
Qy	529	ATGGCTTGGGATATGATGAAGTGGT	556
Db	529		
Qy	601	ATGGCTTGGGATATGATGAAGTGGT	628
Db	601		

Search completed: December 20, 2003, 07:03:12
Job time : 46.7045 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2408.07 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303a-27
Perfect score: 636
Sequence: 1 ATGTTGGTAGGTCATCGA.....ATCACCACCACTATATAG 636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:			
1:	gb_ba:	636	A48689
2:	gb_hgt:	636	AR157339
3:	gb_in:	636	AX452776
4:	gb_om:	636	AX685028
5:	gb_ov:	597.4	AX48687
6:	gb_pat:	597.4	AX157338
7:	gb_ph:	597.4	AX452774
8:	gb_pl:	597.4	AX685026
9:	gb_pr:	597.4	AX48683
10:	gb_ro:	597.4	AR157336
11:	gb_sts:	597.4	AX452772
12:	gb_sy:	515	AX685024
13:	gb_un:	515	AX48667
14:	gb_vi:	515	AR157325
15:	em_ba:	515	AX452754
16:	em_fun:	512	AX685006
17:	em_hum:	512	A48709
18:	em_in:	512	AR157350
19:	em_mu:	512	AX452796
20:	em_om:	512	AX685048
21:	em_or:	512	A48711
22:	em_ov:	512	AR157351
23:	em_pat:	512	AX452798
24:	em_ph:	512	AX685050
25:	em_pl:	459.2	AB008441
26:	em_ro:	459.2	AB008442
27:	em_sts:	459.2	AB008443
28:	em_un:	459.2	AF207766
29:	em_vi:	456	A48669
30:	em_hgt_hum:	456	AR157326
31:	em_hgt_inv:	456	AR157326
32:	em_hgt_other:	456	AR157326
33:	em_hgt_mus:	456	AR157326
34:	em_hgt_pln:	456	AR157326
35:	em_hgt_rod:	456	AR157326
36:	em_hgt_mam:	456	AR157326
37:	em_hgt_vrt:	456	AR157326
38:	em_sy:	456	AR157326
39:	em_hgtgo_hum:	456	AR157326
40:	em_hgtgo_mus:	456	AR157326
41:	em_hgtgo_other:	456	AR157326

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	636	100.0	636	6	A48689	A48689 Sequence 27
2	636	100.0	636	6	AR157339	AR157339 Sequence
3	636	100.0	636	6	AX452776	AX452776 Sequence
4	636	100.0	636	6	AX685028	AX685028 Sequence
5	597.4	93.9	606	6	A48687	A48687 Sequence 25
6	597.4	93.9	606	6	AR157338	AR157338 Sequence
7	597.4	93.9	606	6	AX452774	AX452774 Sequence
8	597.4	93.9	606	6	AX685026	AX685026 Sequence
9	597.4	93.9	723	6	A48683	A48683 Sequence 21
10	597.4	93.9	723	6	AR157336	AR157336 Sequence
11	597.4	93.9	723	6	AX452770	AX452770 Sequence
12	597.4	93.9	723	6	AX685022	AX685022 Sequence
13	556	87.4	561	6	A48685	A48685 Sequence 23
14	556	87.4	561	6	AR157337	AR157337 Sequence
15	556	87.4	561	6	AX452772	AX452772 Sequence
16	556	87.4	561	6	AX685024	AX685024 Sequence
17	515	81.0	795	6	A48667	A48667 Sequence 5
18	515	81.0	795	6	AR157325	AR157325 Sequence
19	515	81.0	795	6	AX452754	AX452754 Sequence
20	515	81.0	795	6	AX685006	AX685006 Sequence
21	512	80.5	2082	6	A48709	A48709 Sequence 47
22	512	80.5	2082	6	AR157350	AR157350 Sequence
23	512	80.5	2082	6	AX452796	AX452796 Sequence
24	512	80.5	2082	6	AX685048	AX685048 Sequence
25	512	80.5	2433	6	A48711	A48711 Sequence 49
26	512	80.5	2433	6	AR157351	AR157351 Sequence
27	512	80.5	2433	6	AX452798	AX452798 Sequence
28	512	80.5	2433	6	AX685050	AX685050 Sequence
29	459.2	72.2	3296	14	AB008441	AB008441 Hepatitis
30	459.2	72.2	3296	14	AB008442	AB008442 Hepatitis
31	459.2	72.2	3296	14	AB008443	AB008443 Hepatitis
32	459.2	72.2	3296	14	AF207766	AF207766 Hepatitis
33	456.4	71.8	633	6	A48669	A48669 Sequence 7
34	456.4	71.8	633	6	AR157326	AR157326 Sequence
35	456.4	71.8	633	6	AX452756	AX452756 Sequence
36	456.4	71.8	633	6	AX685008	AX685008 Sequence
37	456	71.7	9379	14	AF165052	AF165052 Hepatitis
38	456	71.7	9379	14	HCV132996	HCV132996 Hepatitis
39	454.4	71.4	3296	14	AB008445	AB008445 Hepatitis
40	454.4	71.4	3296	14	AB008446	AB008446 Hepatitis
41	454.4	71.4	9344	14	AB049096	AB049096 Hepatitis
42	454.4	71.4	9379	14	HCVPOLYP	HCVPOLYP Hepatitis
43	454.4	71.4	9386	14	AF165056	AF165056 Hepatitis
44	452.8	71.2	1615	14	HPCNS1SPJ	M74813 Hepatitis C
45	452.8	71.2	3296	14	AB008447	AB008447 Hepatitis

ALIGNMENTS

RESULT 1
A48689
LOCUS A48689 Sequence 27 from Patent WO9604385.
DEFINITION A48689
ACCESSION A48689
VERSION A48689.1 GI:2302402
KEYWORDS .
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 636)
AUTHORS Maertens,G., Bosman,F., De.M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 27 15-FEB-1996;

LOCUS	AR157339	636 bp	DNA	linear	PAT 17-OCT-2001
DESCRIPTION	Sequence 27 from patent US 6245503.				

INNOGENETICS NV (BE)	2172273	960215	COMMENT
Other publication CA	2172273	960215	
Other publication AU	3382495	960304	

VERSION	Unknown.
KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unknown.

FEATURES	source	Location/Qualifiers
		1. .636
		/organism="unidentified"
		/molecule_type="genomic DNA"
		/db_xref="taxon:32644"
CDS		1. .633
		/note="unnamed protein product"
		/codon_start=1
		/protein_id="CAA03137.1"
		/db_xref="GI:2302403"
		/translation="MLGKIDTITTCGFADLVGVIPLVGAPLGAARALAHGVRLVDG VNTATGLPNCFSFI FLIALLSCLTVPASAYEVNYSVMYHVNDCSNSIVTEAADM IMHTPGCVCRNNRSNCVVALPTTLAARNASVPTTIRRHVDSQLFTISPRRHETV QNCNCFVCTVCTGMAWMMWNNSTPTLAAVVSOLLRIVIEGRHHHHH"

1. Maertens, G., Bosman, F., De Martynoff, G. and Buysse, M.-A.
 Purified hepatitis C virus envelope proteins for diagnostic and
 therapeutic use
 Patent: US 6245503-A 27 12-JUN-2001;

```

CDS
1..633
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03137.1"
/db_xref="GI:2302403"
/translations="MGKVIDTITCGFADIVGVPIVGPAPLGAARALAHGVTVLEDG
VYATGLNPQCSFI FLALLSLTVPASAYEVNRYGVHVTNDGNSISVVEADAM
IMHTPGCVPRNNRSRNCVWALTPLAARNAASVPTTIRRHVDSQLFTISPRRHETV
QDNCQVFCVTCRSHRSMWMMNNSPTTALVVSOLLRIVIEGRHHHHHH"

```

source	1. 636	/organism="unknown"
BASE COUNT	119 a	203 c
ORIGIN		174 g
		140 t

mat peptide 1. .630
QDCNCSYPGHIITGHRMADMMNNWSPPTTALVVSQLLRIVIEGRHHHHH"
IMETPGVCPVCRNNSSRCWVALPTLAAENASVPTTIRRHVDSQLTISPRRHETV
VNTATGLPCFSIFLALLSLCTVPASAEYVRNVSVMHYTNDSCNSISVVEAAD
/translation= MLENLIDRLCGRFVETV

Query Match	100.0%	Score 636;	DB 6;	Length 636;
Best Local Similarity	100.0%	Pred. No. 6.4e-130;		
Matches 636;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

BASE COUNT	119 a	203 c	174 g	140 t
ORIGIN	/product="unnamed"			

Qy	1	ATGTTGGTAAAGTTCATCGATACCCCTTATATGCGGCTTCGCCGACCTCGTGGGTACATT	60
Qy	1	ATGTTGGTAAAGTTCATCGATACCCCTTATATGCGGCTTCGCCGACCTCGTGGGTACATT	60
Pb	1	ATGTTGGTAAAGTTCATCGATACCCCTTATATGCGGCTTCGCCGACCTCGTGGGTACATT	60

```

Query Match      100.0%; Score 636; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 6.4e-130;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

Qy 1 ATGTTGGGTAAGGTCATCGATACCTTTACATCGCGCTTCGCGGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAGGTCATCGATACCTTTACATCGCGCTTCGCGGACCTCGTGGGTACATT 60

DB	61	121	180
QY	CTGGAGGACGGCGTGAACTATGCAACAGGGAAATTGCCGGTTGCTCTTTCTCTATCTTC	CTGGAGGACGGCGTGAACTATGCAACAGGGAAATTGCCGGTTGCTCTTTCTCTATCTTC	CTGGAGGACGGCGTGAACTATGCAACAGGGAAATTGCCGGTTGCTCTTTCTCTATCTTC

DB	I AIGTGGGAAAGGTCATCGAACCCTACCAAGGCGCCGCGGT	120
QY	61 CCGCTGTCGGCGCCCCCCTTAGGGGGCGCTCCAGAGGCCCTGGCGCATGGCGCTCGGGTT	120
DH	61 CCGCTGTCGGCGCCCCCCTTAGGGGGCGCTGCAGAGGCCCTGGCGCATGGCGCTCGGGTT	120

DB	QY
121	181

DB	61	CGGCTCGTCGGCGCCCGCCCTAGGGGGCGCGCTGCCAGGGGCCCTTGGCCCATAGGGCGTCCGGGTT
QY	121	CTGGAGAGCGGCGTGAATTATGCAACAGGGAATTTGCCGGTTCGTCCTTTCTATATCTTC
+	180	CTGGCGTCGGCGTGAATTATGCAACAGGGAATTTGCCGGTTCGTCCTTTCTATATCTTC

db	181	241	300
db	CTCTTGGCTTTGCTGTCCTCTGTGACCGTTCAGCTTC	CGTTCGCTTCCAGCTTC	CGCTTATGAAGTGGCAACGTG
QY	TCCGGGATGTACCATGTGTACGAA	CGACTGCTCCAACTCAAGCATCTGTATGAGGACGCG	

db	121	181	240
db	CTGGAGGACGGCGTGAACATATGCAACAGGAATTTGCCGGTGGCTCTTCTCTATCTTC		
Qy		CTCTTGGCTTTGCTGTCCTGCTGTGACCGTTCCAGCTTCCGCTTATTAAGTAGTGGCAACGTG	

Db	241	TCCGGGATGTACCAATGTCACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGGCAGCG	300
Qc	301	GACATGATCATGACACACCCCGGGTGGGTGCCCTGGCTTCGGAGAAACAATCTCTTCGGCG	360

	181	241
db	CTCTTTGGCTTTGCTGTCTCTGTCTGACCGTTCCAGCTTCCGCTTATTAAGATGGCGAACGTG	TCGGGGATGACCATGTGTACAGAACGCTGCTCCACTCAAGCATTTGTGTATGAGGCAGCG
QY		

[illegible]

QY	241	QY
241	TCCGGGATGACCATGTTACGAAAGCTGCTCCACTCAAGCATTTGTGTATGAGCAGCG	300
Db		
301	GACATGATCATGCACACCCCGGGTGGCGTGCCTGCGTTTCGGAGAACAACTCTTCCCGC	360
QY		

	361	TGCTGGTAGCGCTACCCCCCAAGCCTCCAGATTCG	420
Qy			
	361	TGCTGGTAGCGCTACCCCCCAAGCCTCCAGATTCG	420
Dd			

Qy 301 GACATGATCATGCACACCCCGGGTGCCTGCGTTTCGGAGAACACTCTTCCCGC 360

Db 301 GACATGATCATGCACACCCCGGGTGCCTGCGTTTCGGAGAACACTCTTCCCGC 360

Qy 361 TGCTGGGCTAGCGCTCAACCCCAAGCTTCGAGGTTAGGAACGCCAGCGTCCCAACCAAGACA 420

	Qy	Db
421	ATACGAGCCGCGT	ATACGAGCCGCGT
422	CGATTCCAGCTGTT	CGATTCCAGCTGTT
423	CACCATCTCGCCT	CACCATCTCGCCT
424	CGCCGCGCATGAG	CGCCGCGCATGAG
425	ACCGG	ACCGG

QY 361 TGCTGGGTAGCGGTCTACCCCGCTCGAGCTTAGGAAAGCCAGAGGTGCCCAACCAACA 424

DB 361 TGCTGGGTAGCGGTCTACCCCGCTCGAGCTTAGGAAAGCCAGAGGTGCCCAACCAACA 420

QY	481	541
QY	CAGGACGCAATTGCTCAATCTATCCGGCCACATACGGGTACCGGTATGGCTTGGAT	541
Dh	CAGGACGCAATTGCTCAATCTATCCGGCCACATACGGGTACCGGTATGGCTTGGAT	541

Qy	421	ATACGAGCCGCACTCGATTCCAGTGTTTCCAATCTGCCTCGCCGGCATGAGACGGTG	480
<hr/>			
Db	421	ATACGAGCCGCACTCGATTCCAGTGTTTCCAATCTGCCTCGCCGGCATGAGACGGTG	480

QY 541 ATGATGATGAACCTGGTGCCTCAACCGCCCTGGTGGTATCGCAGCTGTCCGGATCGTG
600

		540
Qy	481 CAGGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTACCGTATGGCTTGGGAT	540
Dd	481 CAGGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTACCCTATGGCTTGGGAT	540

[illegible]

DB	Qy	D _b
481	541	541
CAGGACATGCGATATCGATGATGATCGAGCTGCTCCGATCGTG	ATGATGATGAACCTGGTGCCTCAACCGCCCTGGTGTATCGAGCTGCTCCGATCGTG	ATGATGATGAACCTGGTGCCTCAACCGCCCTGGTGTATCGAGCTGCTCCGATCGTG

DB 601 ATCGAGGGCAGACACCATCACCACCATCAATAAG 636

DB	541	ATGATGATGAACCTGCTGCCTACAAAGCGCCCTGGTGGTATGCGAGGCTGCCTCCCGGATCGGCGA
Qy <td>601 <td>ATCGAGGGCAGACACCATCACCCACCATCACTAATAG 636</td> </td>	601 <td>ATCGAGGGCAGACACCATCACCCACCATCACTAATAG 636</td>	ATCGAGGGCAGACACCATCACCCACCATCACTAATAG 636
Y <td>501 <td>ATCGATGAGGCAGACACCATCACCCACCATCACTAATAG 636</td> </td>	501 <td>ATCGATGAGGCAGACACCATCACCCACCATCACTAATAG 636</td>	ATCGATGAGGCAGACACCATCACCCACCATCACTAATAG 636

RESU01
AX452776
LOCUS AX452776 DNA linear PAT 06-JUL-20
DEFINITION Sequence 27 from Patent EP1211315.

Db
601 ATCGAGGCAGACCATTCACCCACCATCACTAATAG 636
|||||
RESULT 2

ACCESSION AX452776

RESULT 2
AR157339

QY	421	ATACGACGCCAGTCGATCCAGCTGTTCCAGCTGTTCCAGCTCGCTCGCGGCATGAGACGGTG	480
Db	421	ATACGACGCCAGTCGATCCAGCTGTTCCAGCTGTTCCAGCTCGCTCGCGGCATGAGACGGTG	480
QY	481	CAGGACTCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT	540
Db	481	CAGGACTCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT	540
QY	541	ATGATGATGAATGCTGCGCTACAAACGGCCCTGGTGGTATGCGAGCTGCTCCGGATCGTG	600
Db	541	ATGATGATGAATGCTGCGCTACAAACGGCCCTGGTGGTATGCGAGCTGCTCCGGATCGTG	600
QY	601	ATCGAGGCGAGACACCATCACCACCATCACTAATAG	636
Db	601	ATCGAGGCGAGACACCATCACCACCATCACTAATAG	636
RESULT 5			
LOCUS	A48687	606 bp	DNA linear PAT 07-MAR-1997
DEFINITION	Sequence 25 from Patent WO9604385.		
ACCESSION	A48687		
VERSION	A48687.1	GI:2302400	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 606)		
AUTHORS	Maertens, G., Bosman, F., De M.G. and Buyse, M.		
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE		
JOURNAL	Patent: WO 9604385-A 25 15-FEB-1996;		
COMMENT	INNOGENETICS NV (BE)		
Other publication	CA 2172273 960215		
Other publication	AU 3382495 960304.		
FEATURES	Location/Qualifiers		
source	1..606		
/organism="unidentified"			
/mol_type="genomic DNA"			
/db_xref="taxon:32644"			
1..603			
/note="unnamed protein product"			
/codon_start=1			
/protein_id="CAA03136.1"			
/db_xref="GI:2302401"			
/translation="MLGKVIDTLTCGFADLVGYPLVGAPLGGARALAHGVRLVLEDDVNYATGNLPCSPSIFLLALLSLCTWPASAYEVNVSVMVHTNDCNSISVVEAADMIMHTPGVCPCVRENNSRCSWVLTPLAARNASVPTTIRRHVDSQLFTISPRRHETVQDCNCSYPGHIHTGHRWADMNMNWSPTTALVVSQLLRIL"			
mat_peptide	1..600		
/product="unnamed"			
BASE COUNT	109 a 193 c 167 g 137 t		
ORIGIN			
Query Match	93.9%;	Score 597.4;	DB 6; Length 606;
Best Local Similarity	99.8%;	Pred. No. 2e-121;	
Matches 598;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	ATGTTGGGTAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
QY	61	CGGCTCGTCGGCGCCCTTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTT	120
Db	61	CGGCTCGTCGGCGCCCTTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTT	120
QY	121	CTGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC	180
Db	121	CTGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGCTCTGTCAGCGGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG	240
Db	181	CTCTTGGCTTTGCTGCTCTGTCAGCGGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG	240
QY	241	TCCGGGATGTACATGTGCAAGACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGG	300
Db	241	TCCGGGATGTACATGTGCAAGACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGG	300
QY	301	GACATGATCATGTCACACCCCGGGTCCGCTCGCTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG	360
Db	301	GACATGATCATGTCACACCCCGGGTCCGCTCGCTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG	360
QY	421	ATACGACGCCAGTCGATCCAGCTGTTCCAGCTGTTCCAGCTCGCTCGCGGCATGAGACGGTG	480
Db	421	ATACGACGCCAGTCGATCCAGCTGTTCCAGCTGTTCCAGCTCGCTCGCGGCATGAGACGGTG	480
QY	481	CAGGACTCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT	540
Db	481	CAGGACTCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT	540
QY	541	ATGATGATGAATGCTGCGCTACAAACGGCCCTGGTGGTATGCGAGCTGCTCCGGATCGTG	599
Db	541	ATGATGATGAATGCTGCGCTACAAACGGCCCTGGTGGTATGCGAGCTGCTCCGGATCGTG	599
RESULT 6			
LOCUS	AR157338	606 bp	DNA linear PAT 17-OCT-2001
DEFINITION	Sequence 25 from patent US 6245503.		
ACCESSION	AR157338		
VERSION	AR157338.1	GI:16218271	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 606)		
AUTHORS	Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.		
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use		
JOURNAL	Patent: US 6245503-A 25 12-JUN-2001;		
FEATURES	Location/Qualifiers		
source	1..606		
/organism="unknown"			
BASE COUNT	109 a 193 c 167 g 137 t		
ORIGIN			
Query Match	93.9%;	Score 597.4;	DB 6; Length 606;
Best Local Similarity	99.8%;	Pred. No. 2e-121;	
Matches 598;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	ATGTTGGGTAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
QY	61	CGGCTCGTCGGCGCCCTTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTT	120
Db	61	CGGCTCGTCGGCGCCCTTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTT	120
QY	121	CTGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC	180
Db	121	CTGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGCTCTGTCAGCGGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG	240
Db	181	CTCTTGGCTTTGCTGCTCTGTCAGCGGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG	240
QY	241	TCCGGGATGTACATGTGCAAGACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGG	300
Db	241	TCCGGGATGTACATGTGCAAGACGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGG	300
QY	301	GACATGATCATGTCACACCCCGGGTCCGCTCGCTTCCAGCTTCGGGAGAACCACTCTTCCCGC	360


```

Db      301  GACATGATCATGCACACCCCGGGTGGTGGCTTGGGAGAACAACTCTTCCCGC 360
QY      361  TGCTGGTAGCGCTCACCCCAAGCTCGCAGCTAGGACCCAGCGTCCCAACACGACA 420
Db      361  TGCTGGTAGCGCTCACCCCAAGCTCGCAGCTAGGACCCAGCGTCCCAACACGACA 420
QY      421  ATACGACGCCAGTCCAGTCCAGCTGTTCACCATCTCGCCTCGCGGATGAGACGGTG 480
Db      421  ATACGACGCCAGTCCAGTCCAGCTGTTCACCATCTCGCCTCGCGGATGAGACGGTG 480
QY      481  CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTGGGAT 540
Db      481  CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTGGGAT 540
QY      541  ATGATGATGAATCGTTCGCTACACCGGCCCTGGTGGTATCGCAGCTGCCGGATCGT 599
Db      541  ATGATGATGAATCGTTCGCTACACCGGCCCTGGTGGTATCGCAGCTGCCGGATCGT 599

RESULT 7
AX452774
LOCUS      Hepatitis C virus
DEFINITION Hepatitis C virus
ACCESSION AX452774
VERSION    AX452774.1
KEYWORDS   GI:21712459
SOURCE     Hepatitis C virus
ORGANISM   Hepatitis C virus
REFERENCE  1
AUTHORS    Maertens,G., Bosman,F., de Martynoff,G. and Buysse,M.A.
TITLE      Recombinant vectors for producing hcv envelope proteins
JOURNAL    Patent: EP 1211315-A 25 05-JUN-2002;
            Innogenetics N.V. (BE)
FEATURES   Location/Qualifiers
            source
              1..606
              /organism="Hepatitis C virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:11103"
            CDS
              1..603
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAD38074.1"
              /db_xref="GI:21712460"
              /translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLVDG
              VNYATNLPGCSFSFILALLSLCTVPASAVEVRNVSGMYHVTNDCSSIVYEADM
              IMHFGCVPVCVRENNSSRCWALTFTLAARNASVPTTIRRHVDSQLFTISPRRHETV
              QDCNCSIVPGHITGHRMAWMMNWSPTTALVVSQLLRL"
            mat_peptide
              1..600
            BASE COUNT  109 a 193 c 167 g 137 t
            ORIGIN
              93.9%; Score 597.4; DB 6; Length 606;
              Best Local Similarity 99.8%; Pred. No. 2e-121;
              Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCAGCTCGTGGGGTACATT 60
Db      1  ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCAGCTCGTGGGGTACATT 60
QY      61  CGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGGCCCTGCCAGCTGGCGTCCGGGT 120
Db      61  CGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGGCCCTGCCAGCTGGCGTCCGGGT 120
QY      121  CTGAGGAGCGGCGTGAACCTATGCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTC 180
Db      121  CTGAGGAGCGGCGTGAACCTATGCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTC 180
QY      181  CTCCTGGCTTCTGCTCTGTCTGACCGGTTTCAGCTTCGCTTCGCTTATGAAGTCGCAACG 240

```

```

Db      181  CTCCTGGCTTTGCTGCTCTGTCTGACCGTTTCAGCTTCGGCTTATGAAGTCGCAACG 240
QY      241  TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGG 300
Db      241  TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGG 300
QY      301  GACATGATCATGCACACCCCGGGTGGTGGCTTGGGAGAACAACTCTTCCCGC 360
Db      301  GACATGATCATGCACACCCCGGGTGGTGGCTTGGGAGAACAACTCTTCCCGC 360
QY      361  TGCTGGTAGCGCTCACCCCAAGCTCGCAGCTAGGACCCAGCGTCCCAACACGACA 420
Db      361  TGCTGGTAGCGCTCACCCCAAGCTCGCAGCTAGGACCCAGCGTCCCAACACGACA 420
QY      421  ATACGACGCCAGTCCAGTTCGCTACACCGGCCCTGGTGGTATCGCAGCTGCCGGATG 480
Db      421  ATACGACGCCAGTCCAGTTCGCTACACCGGCCCTGGTGGTATCGCAGCTGCCGGATG 480
QY      481  CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTGGGAT 540
Db      481  CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTGGGAT 540
QY      541  ATGATGATGAATCGTTCGCTACACCGGCCCTGGTGGTATCGCAGCTGCCGGATCGT 599
Db      541  ATGATGATGAATCGTTCGCTACACCGGCCCTGGTGGTATCGCAGCTGCCGGATCGT 599

RESULT 8
AX685026
LOCUS      606 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 25 from Patent WO0205548.
ACCESSION AX685026
VERSION    AX685026.1
KEYWORDS   GI:29371431
SOURCE     Hepatitis C virus
ORGANISM   Hepatitis C virus
REFERENCE  1
AUTHORS    Maertens,G., Bosman,F. and Buysse,M.A.
TITLE      Purified Hepatitis C Virus envelope proteins for diagnostic and
            therapeutic use
JOURNAL    Patent: WO 0205548-A 25 18-JUL-2002;
            INNOGENETICS N.V. (BE)
FEATURES   Location/Qualifiers
            source
              1..606
              /organism="Hepatitis C virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:11103"
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAD86527.1"
              /db_xref="GI:29371432"
              /translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLVDG
              VNYATNLPGCSFSFILALLSLCTVPASAVEVRNVSGMYHVTNDCSSIVYEADM
              IMHFGCVPVCVRENNSSRCWALTFTLAARNASVPTTIRRHVDSQLFTISPRRHETV
              QDCNCSIVPGHITGHRMAWMMNWSPTTALVVSQLLRL"
            mat_peptide
              1..600
            BASE COUNT  109 a 193 c 167 g 137 t
            ORIGIN
              93.9%; Score 597.4; DB 6; Length 606;
              Best Local Similarity 99.8%; Pred. No. 2e-121;
              Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCAGCTCGTGGGGTACATT 60
Db      1  ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCAGCTCGTGGGGTACATT 60
QY      61  CGCTCGTCGGCGCCCGCTAGGGGCGCTGCCAGGGCCCTGCCAGCTGGCGTCCGGGT 120

```


Mon Dec 22 13:28:43 2003

Query Match		93.9%; Score 597; DB 6; Length 723;
Best Local Similarity		100.0%; Pred. No. 2.5e-121; Indels 0; Gaps 0;
Matches 597; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGTTGGGTAAAGGTATCATGATACCTTATACATGCGGCTTCCGACCTCTGTTGGGTACATT 60
Db	1	ATGTTGGGTAAAGGTATCATGATACCTTATACATGCGGCTTCCGACCTCTGTTGGGTACATT 60
QY	61	CGGCTGCTGGGCGCCCTAGGAGGCGCTGCCAGGCGCTTGGGCGCATGGCGTCCGGGTT 120
Db	61	CGGCTGCTGGGCGCCCTAGGAGGCGCTGCCAGGCGCTTGGGCGCATGGCGTCCGGGTT 120
QY	121	CTGGAGGACGGCGTGAACCTATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTC 180
Db	121	CTGGAGGACGGCGTGAACCTATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTC 180
QY	181	CTCTTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db	181	CTCTTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY	241	TCCGGATGTACCATGTCTACGAACGACTGCTCCAATCAAGCATTTGTATGAGGACGG 300
Db	241	TCCGGATGTACCATGTCTACGAACGACTGCTCCAATCAAGCATTTGTATGAGGACGG 300
QY	301	GACATGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db	301	GACATGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY	361	TGCTGGGTAGCGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db	361	TGCTGGGTAGCGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY	421	ATACGACGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db	421	ATACGACGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY	481	CAGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGTACCGGTACCGGTACCGGTACCG 540
Db	481	CAGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGTACCGGTACCGGTACCGGTACCG 540
QY	541	ATGATGATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
Db	541	ATGATGATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
RESULT 10		
AR157336		723 bp DNA linear PAT 17-OCT-2001
LOCUS		Sequence 21 from patent US 6245503.
DEFINITION		AR157336
ACCESSION		AR157336.1 GI:16218269
VERSION		Unknown.
KEYWORDS		Unknown.
SOURCE		Unclassified.
ORGANISM		1 (bases 1 to 723)
REFERENCE		Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
AUTHORS		Purified hepatitis C virus envelope proteins for diagnostic and
TITLE		therapeutic use
JOURNAL		Patent: US 6245503-A 21 12-JUN-2001;
FEATURES		Location/Qualifiers
source		1..723
BASE COUNT		126 a 220 c 208 g 169 t
ORIGIN		/organism="unknown"
Query Match		93.9%; Score 597; DB 6; Length 723;
Best Local Similarity		100.0%; Pred. No. 2.5e-121; Indels 0; Gaps 0;
Matches 597; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGTTGGGTAAAGGTATCATGATACCTTATACATGCGGCTTCCGACCTCTGTTGGGTACATT 60
Db	1	ATGTTGGGTAAAGGTATCATGATACCTTATACATGCGGCTTCCGACCTCTGTTGGGTACATT 60


```
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCTCCGGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCTCCGGGTT 120
QY 121 CTGAGGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
Db 121 CTGAGGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
QY 181 CTCCTTGGCTTTGCTCTGCTGTCTGACCGGTTCCAGCTTCCGCTTATGAAAGTGCACAGCGT 240
Db 181 CTCCTTGGCTTTGCTCTGCTGTCTGACCGGTTCCAGCTTCCGCTTATGAAAGTGCACAGCGT 240
QY 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCG 300
Db 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCGGAGAAACAATCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCGGAGAAACAATCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCCAGCTGCGAGTAGGAACGCCAGCGTCCCCACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCCCAGCTGCGAGTAGGAACGCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCCAGCTGATTCACAGCTGTTCCAGCATCTGCGCTGCGGAGTATGAGACGGTG 480
Db 421 ATACGACGCCAGCTGATTCACAGCTGTTCCAGCATCTGCGCTGCGGAGTATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACGGGTACCGTATGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACGGGTACCGTATGCTTGGGAT 540
QY 541 ATGATGATGAATGGTTCGCTACACGGCCCTGGTGTATCGCAGCTGCTCCGGATC 597
Db 541 ATGATGATGAATGGTTCGCTACACGGCCCTGGTGTATCGCAGCTGCTCCGGATC 597

RESULT 11
AX452770
LOCUS AX452770 723 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 21 from Patent EP1211315.
ACCESSION AX452770
VERSION AX452770.1 GI:21712455
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 21 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/feature="unnamed protein product"
/codon_start=1
/db_xref="GI:21712456"
/translation="MLGKVIDTLTCGFADLVGYIFLVGAPLGGARALAHGVRLVDG
VNYATGNLPGCSFSLFLLALLSLTPASAYEVNVGVHVNDCSSNIVYEADM
INHTPCVPCVRENSSRCWALPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNGSIVPGHTITGRMAWMMNNSPTTALVVSQLLRIPQAVVDMVAGAHGVLAGL
AYISMVGNWAKVLIVMLFAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
ORIGIN
```

```
Query Match 93.9%; Score 597; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.5e-121;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCCGCCAGCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCCGCCAGCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCTCCGGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCTCCGGGTT 120
QY 121 CTGAGGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
Db 121 CTGAGGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
QY 181 CTCCTTGGCTTTGCTCTGCTGTCTGACCGGTTCCAGCTTCCGCTTATGAAAGTGCACAGCGT 240
Db 181 CTCCTTGGCTTTGCTCTGCTGTCTGACCGGTTCCAGCTTCCGCTTATGAAAGTGCACAGCGT 240
QY 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCG 300
Db 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCGGAGAAACAATCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCGGAGAAACAATCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCCAGCTGCGAGTAGGAACGCCAGCGTCCCCACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCCCAGCTGCGAGTAGGAACGCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCCAGCTGATTCACAGCTGTTCCAGCATCTGCGCTGCGGAGTATGAGACGGTG 480
Db 421 ATACGACGCCAGCTGATTCACAGCTGTTCCAGCATCTGCGCTGCGGAGTATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACGGGTACCGTATGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACGGGTACCGTATGCTTGGGAT 540
QY 541 ATGATGATGAATGGTTCGCTACACGGCCCTGGTGTATCGCAGCTGCTCCGGATC 597
Db 541 ATGATGATGAATGGTTCGCTACACGGCCCTGGTGTATCGCAGCTGCTCCGGATC 597

RESULT 12
AX685022
LOCUS AX685022 723 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 21 from Patent WO0205548.
ACCESSION AX685022
VERSION AX685022.1 GI:29371427
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Bosman,F. and Buyse,M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 21 18-JUL-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/feature="unnamed protein product"
/codon_start=1
/db_xref="GI:29371428"
CDS
```


Mon Dec 22 13:28:43 2003

/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRVLEBDG
VNYATGNLPGCSFISFLALLSCLTVPASAEVRNVSGMYHTVNDGNSNLSIVVEAADM
IMHTPGCVPCVRNNSRCWVALTPTLAARNSVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIYPGHITGHRMAWDMNNSPTTALVLSQLLRIFQAVVDVAGAHGVLGAL
AYNSVGNWAKVLIVMLLFAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
ORIGIN

Query Match 93.9%; Score 597; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.5e-121;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGGTATCGATACACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAGGTATCGATACACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGCGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGCGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCCTGTGACCGTTTCCAGCTTTCGGCTTATGAAGTGCACAACGTG 240
DB 181 CTCTTGGCTTTGCTGTCCTGTGACCGTTTCCAGCTTTCGGCTTATGAAGTGCACAACGTG 240
QY 241 TCCGGATGATACCATGTCAAGAACAGTGTCTCAACTCAAGCATTTGTATGAGCGAGG 300
DB 241 TCCGGATGATACCATGTCAAGAACAGTGTCTCAACTCAAGCATTTGTATGAGCGAGG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGTTTCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGTTTCGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGGGTCAACCCACAGCTCGCAGTAGGAACCCAGCGTCCCGACACGACA 420
DB 361 TGCTGGGTAGGGTCAACCCACAGCTCGCAGTAGGAACCCAGCGTCCCGACACGACA 420
QY 421 ATACAGCCACGTCGATTCACGCTTCCACATCTCGCTCGCGGCATGAGCGGTG 480
DB 421 ATACAGCCACGTCGATTCACGCTTCCACATCTCGCTCGCGGCATGAGCGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGGCTGGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGGCTGGGAT 540
QY 541 ATGATGATGAACGTGTCGCTTACAAAGCCGCTGGTGTATCGAGCTGCTCCGGATC 597
DB 541 ATGATGATGAACGTGTCGCTTACAAAGCCGCTGGTGTATCGAGCTGCTCCGGATC 597

RESULT 13
A48685
LOCUS A48685 561 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 23 from Patent WO9604385.
ACCESSION A48685
VERSION A48685.1 GI:2302398
KEYWORDS unidentifed
SOURCE unidentifed
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 561)
AUTHORS Maertens,G., Bosman,F., De M.G. and Buysse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 23 15-FEB-1996;
INNOGENETICS NV (BE)
COMMENT Other publication CA 2172273 960215
Other publication AU 3382495 960304.

FEATURES
source
Location/Qualifiers
1..561
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1..558
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03135.1"
/db_xref="GI:2302398"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRVLEBDG
VNYATGNLPGCSFISFLALLSCLTVPASAEVRNVSGMYHTVNDGNSNLSIVVEAADM
IMHTPGCVPCVRNNSRCWVALTPTLAARNSVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIYPGHITGHRMAWDMNNSPTTALVLSQLLRIFQAVVDVAGAHGVLGAL
AYNSVGNWAKVLIVMLLFAP"

mat_peptide 1..555
/product="unnamed"
BASE COUNT 103 a 176 c 155 g 127 t
ORIGIN

Query Match 87.4%; Score 556; DB 6; Length 561;
Best Local Similarity 100.0%; Pred. No. 2.7e-112;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGGTATCGATACACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAGGTATCGATACACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGCGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGCGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCCTGTGACCGTTTCCAGCTTTCGGCTTATGAAGTGCACAACGTG 240
DB 181 CTCTTGGCTTTGCTGTCCTGTGACCGTTTCCAGCTTTCGGCTTATGAAGTGCACAACGTG 240
QY 241 TCCGGATGATACCATGTCAAGAACAGTGTCTCAACTCAAGCATTTGTATGAGCGAGG 300
DB 241 TCCGGATGATACCATGTCAAGAACAGTGTCTCAACTCAAGCATTTGTATGAGCGAGG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGTTTCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGTTTCGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGGGTCAACCCACAGCTCGCAGTAGGAACCCAGCGTCCCGACACGACA 420
DB 361 TGCTGGGTAGGGTCAACCCACAGCTCGCAGTAGGAACCCAGCGTCCCGACACGACA 420
QY 421 ATACAGCCACGTCGATTCACGCTTCCACATCTCGCTCGCGGCATGAGCGGTG 480
DB 421 ATACAGCCACGTCGATTCACGCTTCCACATCTCGCTCGCGGCATGAGCGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGGCTGGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGGCTGGGAT 540
QY 541 ATGATGATGAACGTGTT 556
DB 541 ATGATGATGAACGTGTT 556

RESULT 14
AR157337
LOCUS AR157337 561 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 23 from patent US 6245503.
ACCESSION AR157337
VERSION AR157337.1 GI:16218270
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 175.828 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303a-27

Perfect score: 636

Sequence: 1 ATGTTGGTAGGTCATCGA.....ATCACCACCATCACTAATAG 636

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	636	100.0	636	17	AAT12964 HCV E1 construct H
2	636	100.0	636	24	AAL48928 Hepatitis C virus
3	597.4	93.9	606	17	AAT12963 HCV E1 construct H
4	597.4	93.9	606	24	AAL48927 Hepatitis C virus
5	597	93.9	723	17	AAT12961 HCV E1 construct H
6	597	93.9	723	24	AAL48925 Hepatitis C virus
7	556	87.4	561	17	AAT12962 HCV E1 construct H
8	556	87.4	561	24	AAL48926 Hepatitis C virus

9	515	81.0	795	17	AAT12705 HCV E1 construct H
10	515	81.0	795	24	AAL48914 Hepatitis C virus
11	512	80.5	2082	24	AAL48939 Hepatitis C virus
12	512	80.5	2086	17	AAT12973 HCV E1 construct H
13	512	80.5	2433	17	AAT12974 Hepatitis C virus
14	501	78.8	2434	24	AAL48940 HCV E1 construct H
15	456.4	71.8	633	17	AAT12706 HCV E1 construct H
16	456.4	71.8	633	24	AAL48915 Hepatitis C virus
17	451.6	71.0	636	17	AAT12709 HCV E1 construct H
18	451.6	71.0	636	24	AAL48918 Hepatitis C virus
19	451.4	71.0	673	19	AAV42305 Cuticle protein c
20	451.2	70.9	2187	19	ABA03491 NANB hepatitis vir
21	451.2	70.9	2340	14	AQ043889 NANBH genomic fra
22	451.2	70.9	2540	15	AAQ63753 Hepatitis C virus
23	451.2	70.9	9605	24	ABK91411 Hepatitis C virus
24	451.2	70.9	9605	24	ABK91424 Hepatitis C virus
25	451.2	70.9	9605	24	ABK91425 Hepatitis C virus
26	451.2	70.9	9605	24	ABK91426 Hepatitis C virus
27	451.2	70.9	9605	24	ABK91428 Hepatitis C virus
28	451.2	70.9	9605	24	ABK91429 Hepatitis C virus
29	451.2	70.9	9605	24	ABK91430 Hepatitis C virus
30	451.2	70.9	9605	24	ABK91431 Hepatitis C virus
31	451.2	70.9	9605	24	ABK91432 Hepatitis C virus
32	451.2	70.9	9605	24	ABK91433 Hepatitis C virus
33	451.2	70.9	9605	24	AAQ25332 Hepatitis C virus
34	451.2	70.9	9608	24	ABK91427 Hepatitis C virus
35	451.2	70.9	11062	24	AAQ25331 Hepatitis C virus
36	451.2	70.9	11076	21	AAA98965 Hepatitis C virus
37	449.6	70.7	1562	19	AAV60672 Fragment #5 isolat
38	449.6	70.7	1880	13	AAQ24467 NANB hepatitis vir
39	449.6	70.7	1953	25	AAQ55222 Plasmid pIDK2 DNA
40	449.6	70.7	2829	19	AAV60673 Fragment #6 isolat
41	448	70.4	1251	13	AAQ26981 HCV gene 1. Hepat
42	448	70.4	2540	13	AAQ29628 Hepatitis C virus
43	448	70.4	3360	17	AAT03677 Hepatitis C genome
44	448	70.4	3461	15	AAQ64068 Non-A, non-B hepat
45	448	70.4	3461	16	AAT30386 5'UTR/CORE/ENV/NS1

ALIGNMENTS

```

RESULT 1
AAT12964
ID AAT12964 standard; DNA; 636 BP.
XX
AC AAT12964;
XX
DT 24-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC140.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
(PINNO-) INNOGENETICS NV.
XX
Bosman F, Buyse M, De Martynoff G, Maertens G;
WPI; 1996-129401/13.
Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

```


PT proteins - in presence of di-sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV

PS Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.

XX Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;

Query Match 100.0%; Score 636; DB 17; Length 636;
Best Local Similarity 100.0%; Pred. No. 1.6e-158;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCTGGGTAAGTCATCGATACCCCTTATCATGGGGTTCGCGACCTCGTGGGTACATT 60
DB 1 ATGTGGGTAAGTCATCGATACCCCTTATCATGGGGTTCGCGACCTCGTGGGTACATT 60
QY 61 CCGCTGTCGGCGCCCTAGGGGGCTGCCAGGGCTCTGGCGATCGCGTCCGGGT 120
DB 61 CCGCTGTCGGCGCCCTAGGGGGCTGCCAGGGCTCTGGCGATCGCGTCCGGGT 120
QY 121 CTGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTCCTTTCTATCTTC 180
DB 121 CTGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTCCTTTCTATCTTC 180
QY 181 CTCTTGGCTTGTCTGTCGTGACCGTTCAGCTTCCGCTTATGAAGTGCCCAACGTG 240
DB 181 CTCTTGGCTTGTCTGTCGTGACCGTTCAGCTTCCGCTTATGAAGTGCCCAACGTG 240
QY 241 TCCGGGATGTACCATGTGACGACGACTGTCTCAACTCAAGCAATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACCATGTGACGACGACTGTCTCAACTCAAGCAATTTGTATGAGGACGG 300
QY 301 GACATGATCATGCAACACCCCGGGTGTGCGTTCGCTTGGGAGAACACTCTTCCCG 360
DB 301 GACATGATCATGCAACACCCCGGGTGTGCGTTCGCTTGGGAGAACACTCTTCCCG 360
QY 361 TGCTGGGTAGCGCTACCCCGACCTCGAGTACAGGACGCGTCCCGACACGACA 420
DB 361 TGCTGGGTAGCGCTACCCCGACCTCGAGTACAGGACGCGTCCCGACACGACA 420
QY 421 ATACGACCCACGTCGATTTCCAGCTGTTACCACTCTCGCTCCCGGATGAGACGGTG 480
DB 421 ATACGACCCACGTCGATTTCCAGCTGTTACCACTCTCGCTCCCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATGTCTCAATATATCCCGCCACATACGGGTACCGTATGGCTTGGAT 540
DB 481 CAGGACTGCAATGTCTCAATATATCCCGCCACATACGGGTACCGTATGGCTTGGAT 540
QY 541 ATGATGATGAATGTCTGCTACACGCGCTTGTGGTATCGAGTCTCTCCGATCGTG 600
DB 541 ATGATGATGAATGTCTGCTACACGCGCTTGTGGTATCGAGTCTCTCCGATCGTG 600
QY 601 ATCGAGGGCAGACACCATCACCACCATCACTAATAG 636

Db 601 ATCGAGGGCAGACACCATCACCACCATCACTAATAG 636

RESULT 2

AAL48928

ID AAL48928 standard; DNA; 636 BP.

XX AAL48928,

AC 24-OCT-2002 (first entry)

XX

DE Hepatitis C virus clone HCC140 E1 protein coding sequence.

XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;

KW virucide; immunostimulant; vaccine; ds.

XX

OS Hepatitis C virus.

XX

FN W0200255548-A2.

XX

PD 18-JUL-2002.

XX

XX 11-JAN-2002; 2002WO-EP00219.

XX

XX 11-JAN-2001; 2001US-260699P.

XX

XX 30-AUG-2001; 2001US-315768P.

XX

XX (INNO-) INNOGENETICS NV.

PA

XX Maertens G, Bosman F, Buyse M;

XX

XX WPI; 2002-599657/64.

XX

XX P-PSDB; AAO18669.

XX

PT New therapeutic vaccine compositions comprising at least one purified

PT recombinant hepatitis C virus (HCV) single or specific oligomeric

PT recombinant envelope protein E1 or E2, useful for immunizing humans

PT from HCV infection

XX

XX Example 2; Page 179-180; 243pp; English.

XX

XX The present invention relates to new therapeutic vaccine compositions for

CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a

CC composition containing at least one purified recombinant HCV single or

CC specific oligomeric recombinant envelope proteins selected from an E1 and

CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are

CC useful for inducing HCV-specific antibodies or for immunising humans

CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as

CC vaccines or therapeutics, in HCV screening and confirmatory antibody

CC tests, for raising antibodies, in the preparation of medicament, and for

CC in vitro monitoring of HCV disease or prognosing the response to

CC treatment of patients suffering from HCV infection. The present sequence

CC is a coding sequence described in the exemplification of the invention.

XX Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;

Query Match 100.0%; Score 636; DB 24; Length 636;

Best Local Similarity 100.0%; Pred. No. 1.6e-158; Indels 0; Gaps 0;

Matches 636; Conservative 0; Mismatches 0;

QY 1 ATGTTGGTAAAGTCATCGATACCCCTTATCATGGGGTTCGCGACCTCGTGGGTACATT 60

DB 1 ATGTTGGTAAAGTCATCGATACCCCTTATCATGGGGTTCGCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTCCGGCGCCCTAGGGGGCTGCGAGGCCCTGGCGATGGCGTCCGGGT 120

DB 61 CCGCTCGTCCGGCGCCCTAGGGGGCTGCGAGGCCCTGGCGATGGCGTCCGGGT 120

QY 121 CTGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTCCTTTCTATCTTC 180

DB 121 CTGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTCCTTTCTATCTTC 180

QY 181 CTCTTGGCTTGTCTGTCGTCTGACCGGTTCCAGCTTCGCTATGAAAGTGCGCAACGTG 240

virucide; immunostimulant; vaccine; ds.

Hepatitis C virus.

WO200255548-A2.

18-JUL-2002.

11-JAN-2002; 2002WO-EP00219.

11-JAN-2001; 2001US-260699P.

30-AUG-2001; 2001US-315768P.

(INNO-) INNOGENETICS NV.

Maertens G, Bosman F, Buyse M;

WPI; 2002-599657/64.

P-PSDB; AA018668.

New therapeutic vaccine compositions comprising at least one purified recombinant hepatitis C virus (HCV) single or specific oligomeric recombinant envelope protein E1 or E2, useful for immunizing humans from HCV infection -

Example 2; Page 177-178; 243pp; English.

The present invention relates to new therapeutic vaccine compositions for inducing hepatitis C virus (HCV)-specific antibodies, comprising a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope protein selected from an E1 and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are useful for inducing HCV-specific antibodies or for immunizing humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, for raising antibodies, in the preparation of medicament, and for in vitro monitoring of HCV disease or prognosis the response to treatment of patients suffering from HCV infection. The present sequence is a coding sequence described in the exemplification of the invention.

Sequence 606 BP; 109 A; 193 C; 167 G; 137 T; 0 other;

Query Match 93.9%; Score 597.4; DB 24; Length 606;

Best Local Similarity 99.8%; Pred. No. 2.6e-148; Indels 0; Gaps 0;

Matches 598; Conservative 0; Mismatches 1;

1 ATGTGGGTAAAGTCAATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60

1 ATGTGGGTAAAGTCAATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60

61 CCGCTCGTGGCGGCCCCCTTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGTT 120

61 CCGCTCGTGGCGGCCCCCTTAGGGGCGCTCCAGGGCCCTGGCGCATGGCTCCGGTT 120

121 CTGGAGGCGGGTGAATGATGCAACAGGGAATTGCCCGGTTCTCTTCTTCTATCTTC 180

121 CTGGAGGCGGGTGAATGATGCAACAGGGAATTGCCCGGTTCTCTTCTTCTATCTTC 180

181 CTCTTGGCTTTGCTCTCTCTGCTGACCGTTCCAGTTCCTGATAGTGGCGCAAGTG 240

181 CTCTTGGCTTTGCTCTCTCTGCTGACCGTTCCAGTTCCTGATAGTGGCGCAAGTG 240

241 TCCGGATGTACATGTACGACGACATGCTCCAACTCAAGCATTTGTATGAGGCAGCG 300

241 TCCGGATGTACATGTACGACGACATGCTCCAACTCAAGCATTTGTATGAGGCAGCG 300

301 GACATGATCATGACACACCCCGGGTGGCTCCCTCGTTCGGGAGAAACAATCTTCCGCG 360

301 GACATGATCATGACACACCCCGGGTGGCTCCCTCGTTCGGGAGAAACAATCTTCCGCG 360

361 TGTGGGTAGCGTCAACCCCGAGCTCGCAGTAGGAGCGCCAGCGTCCCAACACGACA 420

361 TGTGGGTAGCGTCAACCCCGAGCTCGCAGTAGGAGCGCCAGCGTCCCAACACGACA 420

QY 421 ATACGACGCCAGTCGATTCAGCTGTTCCAGCTGTTCCACCATCTCGCTCGCGCATGAGACGGTG 480
DB 421 ATACGACGCCAGTCGATTCAGCTGTTCCAGCTGTTCCACCATCTCGCTCGCGCATGAGACGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTTCACCGTATGGCTTGGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACCTGCTCGCTCAACAGCGCCCTGGTGGTATCGAGCTGCTCCGGATCCT 599
DB 541 ATGATGATGAACCTGCTCGCTCAACAGCGCCCTGGTGGTATCGAGCTGCTCCGGATCCT 599

RESULT 5

AAT12961

ID AAT12961 standard; DNA; 723 BP.

XX AC AAT12961;

XX 24-SEP-1996 (first entry)

XX HCV E1 construct HCC137.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;

XX Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP03031.

XX 29-JUL-1994; 94EP-0870132.

XX (INNO-) INNOGENETICS NV.

XX Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - in presence of di-sulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

XX Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified with a carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

XX Sequence 723 BP; 126 A; 220 C; 208 G; 169 T; 0 other;

XX Query Match 93.9%; Score 597; DB 17; Length 723;

Best Local Similarity 100.0%; Pred. No. 3.5e-148;		Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGTTGGGTAAGGTATCATGATACCCCTTACATCGCGCTTCGCCAGCTCTCGGGGTACATT	60
Db	1	ATGTTGGGTAAGGTATCATGATACCCCTTACATCGCGCTTCGCCAGCTCTCGGGGTACATT	60
QY	61	CGCTCGTTCGGCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT	120
Db	61	CGCTCGTTCGGCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT	120
QY	121	CTGGAGGAGCGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC	180
Db	121	CTGGAGGAGCGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC	180
QY	181	CTCTGGCTTTGCTGTCCTGTGACCGTTCAGCTTCGCTTATGAAGTGCGCAACGTG	240
Db	181	CTCTGGCTTTGCTGTCCTGTGACCGTTCAGCTTCGCTTATGAAGTGCGCAACGTG	240
QY	241	TCGGGATGATCATGTCAGCAACCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT	300
Db	241	TCGGGATGATCATGTCAGCAACCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT	300
QY	361	TCGCTGCTCGCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT	420
Db	361	TCGCTGCTCGCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT	420
QY	421	ATACGAGCGCAGTGCATTCAGCTGTTCACATCTCGCTCGCGGATGAGACGGTG	480
Db	421	ATACGAGCGCAGTGCATTCAGCTGTTCACATCTCGCTCGCGGATGAGACGGTG	480
QY	481	CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGTATGGCTCGGAT	540
Db	481	CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGTATGGCTCGGAT	540
QY	541	ATGATGATGAACGTGTCGCTACAGCGCCCTGGTGGTATCGCAGCTGCTCGGATC	597
Db	541	ATGATGATGAACGTGTCGCTACAGCGCCCTGGTGGTATCGCAGCTGCTCGGATC	597
RESULT 6			
AA148925	AA148925 standard; DNA; 723 BP.		
AC	AA148925;		
DT	24-OCT-2002 (first entry)		
DE	Hepatitis C virus clone HC137 E1 protein coding sequence.		
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;		
OS	vaccine; immunostimulant; vaccine; ds.		
PN	Hepatitis C virus.		
PD	WO200255548-A2.		
PF	18-JUL-2002.		
PR	11-JAN-2002; 2002WO-BP00219.		
PR	11-JAN-2001; 2001US-260699P.		
PR	30-AUG-2001; 2001US-315768P.		
PA	(INNO-) INNOGENETICS NV.		
PI	Maertens G, Bosman F, Buyse M;		
DR	WPI; 2002-599657/64.		

DR	P-PSDB; AA018666.
PT	New therapeutic vaccine compositions comprising at least one purified recombinant hepatitis C virus (HCV) single or specific oligomeric protein E1 or E2, useful for immunizing humans from HCV infection
PT	Example 2; Page 173-174; 243pp; English.
XX	The present invention relates to new therapeutic vaccine compositions for inducing hepatitis C virus (HCV)-specific antibodies, comprising a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope proteins selected from an E1 and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are useful for inducing HCV-specific antibodies or for immunizing humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, for raising antibodies, in the preparation of medicament, and for in vitro monitoring of HCV disease or prognosing the response to treatment of patients suffering from HCV infection. The present sequence is a coding sequence described in the exemplification of the invention.
XX	Sequence 723 BP; 126 A; 220 C; 208 G; 169 T; 0 other;
Query Match 93.9%; Score 597; DB 24; Length 723;	
Best Local Similarity 100.0%; Pred. No. 3.5e-148;	
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGTTGGGTAAGGTATCATGATACCCCTTACATCGCGCTTCGCCAGCTCTCGGGGTACATT 60
Db	1 ATGTTGGGTAAGGTATCATGATACCCCTTACATCGCGCTTCGCCAGCTCTCGGGGTACATT 60
QY	61 CGCTCGTTCGGCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT 120
Db	61 CGCTCGTTCGGCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT 120
QY	121 CTGGAGGAGCGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
Db	121 CTGGAGGAGCGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
QY	181 CTCTGGCTTTGCTGTCCTGTGACCGTTCAGCTTCGCTTATGAAGTGCGCAACGTG 240
Db	181 CTCTGGCTTTGCTGTCCTGTGACCGTTCAGCTTCGCTTATGAAGTGCGCAACGTG 240
QY	241 TCGGGATGATCATGTCAGCAACCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT 300
Db	241 TCGGGATGATCATGTCAGCAACCGCCCTAGGGGCGCTGCAGGGCCCTGCGATGGCTCGGGT 300
QY	301 GACATGATCATGACACACCCCGGGTGCCTGCGTTCGGGAGAACAACTCTTCCCGC 360
Db	301 GACATGATCATGACACACCCCGGGTGCCTGCGTTCGGGAGAACAACTCTTCCCGC 360
QY	361 TCGTGGTGGCTCACCCCGCGCTCGAGTTCAGGAGCCAGCGTCCCGCCACGACA 420
Db	361 TCGTGGTGGCTCACCCCGCGCTCGAGTTCAGGAGCCAGCGTCCCGCCACGACA 420
QY	421 ATACGAGCGCAGTGCATTCAGCTGTTCACATCTCGCTCGCGGATGAGACGGTG 480
Db	421 ATACGAGCGCAGTGCATTCAGCTGTTCACATCTCGCTCGCGGATGAGACGGTG 480
QY	481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGTATGGCTCGGAT 540
Db	481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGTATGGCTCGGAT 540
QY	541 ATGATGATGAACGTGTCGCTACAGCGCCCTGGTGGTATCGCAGCTGCTCGGATC 597
Db	541 ATGATGATGAACGTGTCGCTACAGCGCCCTGGTGGTATCGCAGCTGCTCGGATC 597
RESULT 7	
AA12962	
ID	AA12962 standard; DNA; 561 BP.
XX	

Mon Dec 22 13:28:43 2003

```

AC AAT12962;
XX 24-SEP-1996 (first entry)
XX HCV E1 construct HCV138.
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX ss.
XX Hepatitis C virus.
XX WO9604385-A2.
XX 31-JUL-1995; 95WO-BP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
XX proteins - in presence of disulphide bond cleavage agent, to
XX produce proteins suitable for direct use in vaccines or diagnostic
XX assays of HCV
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or
XX the genotype of HCV. The new purification method preserves the
XX conformation of the recombinantly expressed E1, E2 and E1/E2, and
XX eliminates contaminating proteins. Antigens isolated using this method
XX are more reactive with human sera than those isolated by known
XX techniques.
XX Sequence 561 BP; 103 A; 176 C; 155 G; 127 T; 0 other;
SQ
Query Match 87.4%; Score 556; DB 17; Length 561;
Best Local Similarity 100.0%; Pred. No. 2.3e-137; Indels 0; Gaps 0;
Matches 556; Conservative 0; Mismatches 0;
QY 1 ATGTTGGTGAAGTATCATGATACCTTATATGCGGCTTCGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTGAAGTATCATGATACCTTATATGCGGCTTCGCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGCGCCCTTACGAGGCGCTGCGAGGCGCTGCGCATGGCGTTCGGGTT 120
DB 61 CCGCTCGTGGCGCCCTTACGAGGCGCTGCGAGGCGCTGCGCATGGCGTTCGGGTT 120
QY 121 CTGAGGAGCGGTGAATATGCAACAGGCAATTTGCCGGTGTCTTTCTTATCTTC 180
DB 121 CTGAGGAGCGGTGAATATGCAACAGGCAATTTGCCGGTGTCTTTCTTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGTCTGATGACCGTTCCAGCTTCGCGTATGAGTGGCAACGTG 240
DB 181 CTCTTGGCTTTGCTGCTGTCTGATGACCGTTCCAGCTTCGCGTATGAGTGGCAACGTG 240

```

RESULT 8

```

AAL48926
ID AAL48926 standard; DNA; 561 BP.
XX AC AAL48926;
XX DT 24-OCT-2002 (first entry)
XX DE Hepatitis C virus clone HCV138 E1 protein coding sequence.
XX KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX virucide; immunostimulant; vaccine; ds.
XX OS Hepatitis C virus.
XX PN WO20025548-A2.
XX PD 18-JUL-2002.
XX PF 11-JAN-2002; 2002WO-EP00219.
XX PR 11-JAN-2001; 2001US-260699P.
XX PR 30-AUG-2001; 2001US-315768P.
XX PA (INNO-) INNOGENETICS NV.
XX PI Maertens G, Bosman F, Buyse M;
XX DR WPI; 2002-599657/64.
XX P-PSDB; AAO18667.
XX New therapeutic vaccine compositions comprising at least one purified
XX recombinant hepatitis C virus (HCV) single or specific oligomeric
XX recombinant envelope protein E1 or E2, useful for immunizing humans
XX from HCV infection.
XX Example 2; Page 175-176; 243pp; English.
XX The present invention relates to new therapeutic vaccine compositions for
XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
XX composition containing at least one purified recombinant HCV single or
XX specific oligomeric recombinant envelope proteins selected from an E1 and
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
XX useful for inducing HCV-specific antibodies or for immunising humans
XX against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
XX vaccines or therapeutics, in HCV screening and confirmatory antibody
XX tests, for raising antibodies, in the preparation of medicament, and for
XX

```


CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 561 BP; 103 A; 176 C; 155 G; 127 T; 0 other;

Query Match 87.4%; Score 556; DB 24; Length 561;
Best Local Similarity 100.0%; Pred. No. 2.3e-137; Indels 0; Gaps 0;
Matches 556; Conservative 0; Mismatches 0;
Qy 1 ATGTTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60
Qy 61 CCGCTCGTGGGCGCCCTTAGGGGCGCTGCGAGGCGCTGCGCATGGGTCGCGGTT 120
Db 61 CCGCTCGTGGGCGCCCTTAGGGGCGCTGCGAGGCGCTGCGCATGGGTCGCGGTT 120
Qy 121 CTGGAGGACGGCGTGAATATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGCGTGAATATGCAACAGGAATTTGCCGGTGTCTTCTCTATCTTC 180
Qy 181 CTCCTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCGCGTTATGAAGTGCACACG 240
Db 181 CTCCTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCGCGTTATGAAGTGCACACG 240
Qy 241 TCCGGGATGATACCTGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGC 300
Db 241 TCCGGGATGATACCTGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGC 300
Qy 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGTTCGGGAGAACACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGTTCGGGAGAACACTCTTCCCGC 360
Qy 361 TGCTGGGTAGCGCTACCCCGGCGCTGCGAGTACGAGCGCCAGCGTCCCAACACGACA 420
Db 361 TGCTGGGTAGCGCTACCCCGGCGCTGCGAGTACGAGCGCCAGCGTCCCAACACGACA 420
Qy 421 ATAGGACCCACGTCGATTCACCGCTGTTCCACCTCTCGCCTCGCGCATGAGACGGTG 480
Db 421 ATAGGACCCACGTCGATTCACCGCTGTTCCACCTCTCGCCTCGCGCATGAGACGGTG 480
Qy 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAACGGGTACCGTATGGCTTGGGAT 540
Qy 541 ATGATGATGAATGGT 556
Db 541 ATGATGATGAATGGT 556

RESULT 9
AAT12705
ID AAT12705 standard; DNA; 795 BP.
XX
AC AAT12705;
XX
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC110A.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW 88.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX

PR 29-JUL-1994; 94EP-0870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
XX WPI; 1996-129401/13.
XX
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of disulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
XX techniques.
XX
SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match 81.0%; Score 515; DB 17; Length 795;
Best Local Similarity 89.2%; Pred. No. 1.8e-126;
Matches 597; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
Qy 1 ATGTTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT 60
Qy 61 CCGCTCGTGGGCGCCCTTAGGGGCGCTGCGAGGCGCTGCGCATGGCGTTCGCGGTT 120
Db 61 CCGCTCGTGGGCGCCCTTAGGGGCGCTGCGAGGCGCTGCGCATGGCGTTCGCGGTT 120
Qy 121 CTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
Qy 181 CTCCTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCGTTATGAAGTGCACACG 240
Db 181 CTCCTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCGTTATGAAGTGCACACG 240
Qy 241 TCCGGGATGATACCTGTCACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGCAGC 300
Db 241 TCCGGGATGATACCTGTCACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGCAGC 300
Qy 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGTTCGGGAGAACACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGTTCGGGAGAACACTCTTCCCGC 360
Qy 361 TGCTGGGTAGCGCTACCCCGGCGCTGCGAGTACGAGCGCCAGCGTCCCAACACGACA 420
Db 361 TGCTGGGTAGCGCTACCCCGGCGCTGCGAGTACGAGCGCCAGCGTCCCAACACGACA 420
Qy 421 ATAGGACCCACGTCGATTCACCGCTGTTCCACCTCTCGCCTCGCGCATGAGACGGTG 438
Db 421 ATAGGACCCACGTCGATTCACCGCTGTTCCACCTCTCGCCTCGCGCATGAGACGGTG 480
Qy 439 -----TCCAGCTGTTTCCACCATCTCGCCTCGCGCG 468

Db	481	GGGACCTCTGCGGATCTGCTTCCTCGCTCCCGAGCTGTTCAACCATCTCGCCTCGCGG	540
Qy	469	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGGTCACCGT	528
Db	541	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGGTCACCGT	600
Qy	529	ATGGCTTTGGGATATGATGAACTGGTGCCTCTACCAACCGCCCTCGTGGTATCGCAGCTG	588
Db	601	ATGGCTTTGGGATATGATGAACTGGTGCCTCTACCAACCGCCCTCGTGGTATCGCAGCTG	660
Qy	589	CTCCGGATC 597	
Db	661	CTCCGGATC 669	
RESULT 10			
ID	AAL48914		
XX	AAL48914 standard; DNA; 795 BP.		
AC	AAL48914;		
XX	24-OCT-2002 (first entry)		
DT	Hepatitis C virus clone HCC110A E1 protein coding sequence.		
DE	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;		
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;		
KW	virucide; immunostimulant; vaccine; ds.		
XX	Hepatitis C virus.		
OS	Hepatitis C virus.		
XX	WO200255548-A2.		
PN	18-JUL-2002.		
XX	11-JAN-2002; 2002WO-EP00219.		
PF	11-JAN-2001; 2001US-260699P.		
PR	30-AUG-2001; 2001US-315768P.		
XX	(INNO-) INNOGENETICS NV.		
XX	Maertens G, Bosman F, Buyse M;		
PI	WPI; 2002-599657/64.		
DR	P-PSDB; AAO18661.		
XX	New therapeutic vaccine compositions comprising at least one purified		
XX	recombinant hepatitis C virus (HCV) single or specific oligomeric		
PT	recombinant envelope protein E1 or E2, useful for immunizing humans		
PT	from HCV infection		
XX	Example 2; Page 161-162; 243pp; English.		
PS	The present invention relates to new therapeutic vaccine compositions for		
XX	inducing hepatitis C virus (HCV)-specific antibodies, comprising a		
CC	composition containing at least one purified recombinant HCV single or		
CC	specific oligomeric recombinant envelope protein selected from an E1 and		
CC	an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are		
CC	useful for inducing HCV-specific antibodies or for immunising humans		
CC	against HCV. The recombinant HCV E1 and/or E2 proteins are useful as		
CC	vaccines or therapeutics, in HCV screening and confirmatory antibody		
CC	tests, for raising antibodies, in the preparation of medicament, and for		
CC	in vitro monitoring of HCV disease or prognosing the response to		
CC	treatment of patients suffering from HCV infection. The present sequence		
CC	is a coding sequence described in the exemplification of the invention.		
XX	Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;		
XX	Query Match 81.0%; Score 515; DB 24; Length 795;		
XX	Best Local Similarity 89.2%; Pred. No. 1.8e-126; Indels 72; Gaps 1;		
XX	Matches 597; Conservative 0; Mismatches 0;		

Db 4 TTGGTAAAGTATCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 63
QY 64 CTCGTCGCGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGGGCATGGGTCCGGTTCTG 123
Db 64 CTCGTCGCGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGGGCATGGGTCCGGTTCTG 123
QY 124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTGCTCTTCTCTATCTTCCTC 183
Db 124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTGCTCTTCTCTATCTTCCTC 183
QY 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGAAGTGTGCC 243
Db 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGAAGTGTGCC 243
QY 244 GGGATGTACATGTCAACGAACTGCTCAACTCAAGCATTTGTATGAGGCGGCGAC 303
Db 244 GGGATGTACATGTCAACGAACTGCTCAACTCAAGCATTTGTATGAGGCGGCGAC 303
QY 304 ATGATCATGACACACCCCGGGTGGTGGCTTCGCTTCCGAGAAACAATCTTCCGCTGC 363
Db 304 ATGATCATGACACACCCCGGGTGGTGGCTTCGCTTCCGAGAAACAATCTTCCGCTGC 363
QY 364 TGGGTAGCGCTCACCCCGGCGCTGCGAGTGGAGACGCGCTCCCAACAGCAATA 423
Db 364 TGGGTAGCGCTCACCCCGGCGCTGCGAGTGGAGACGCGCTCCCAACAGCAATA 423
QY 424 CGACGCCACGTGCAT----- 438
Db 424 CGACGCCACGTGCATTTGCTGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGG 483
QY 439 -----TCCAGCTTTACCATCTCGCTCGCGGAT 471
Db 484 GACCTCTGCGGATCTGCTTCTCCGCTCCAGCTGTTCAACATCTCGCTCGCGGAT 543
QY 472 GAGACGGTGCAGACTGCAATTTGCTCAATCTATCCCGGCAATATACGGGTACCGTATG 531
Db 544 GAGACGGTGCAGACTGCAATTTGCTCAATCTATCCCGGCAATATACGGGTACCGTATG 603
QY 532 GCTTGGGATATGATGAATGCTGCTGCTTACCAACGCGCTGCTGATGCGAGCTGCTC 591
Db 604 GCTTGGGATATGATGAATGCTGCTGCTTACCAACGCGCTGCTGATGCGAGCTGCTC 663
QY 592 CGGATC 597
Db 664 CGGATC 669

RESULT 13
ID AAT12974
XX AAT12974 standard; DNA; 2433 BP.
XX AC AAT12974;
XX DT 25-SEP-1996 (first entry)
XX HCV E1 construct HCC166.
XX DE HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX OS ss.
XX OS Hepatitis C virus.
XX PN WO9604385-A2.
XX PD 15-FEB-1996.
XX PF 31-JUL-1995; 95WO-EP03031.
XX PR 29-JUL-1994; 94EP-0870132.
XX PA (INNO-) INNOGENETICS NV.
XX

PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
XX proteins - in presence of disulphide bond cleavage agent, to
XX produce proteins suitable for direct use in vaccines or diagnostic
XX assays of HCV
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or
XX the genotype of HCV. The new purification method preserves the
XX conformation of the recombinantly expressed E1, E2 and E1/E2, and
XX eliminates contaminating proteins. Antigens isolated using this method
XX are more reactive with human sera than those isolated by known
XX techniques.
XX Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;
SQ
Query Match 80.5%; Score 512; DB 17; Length 2433;
Best Local Similarity 89.2%; Pred. No. 1.5e-125;
Matches 594; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 4 TTGGTAAAGTATCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 63
Db 355 TTGGTAAAGTATCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCGG 414
QY 64 CTCGTCGCGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGGGCATGGGTCCGGTTCTG 123
Db 415 CTCGTCGCGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGGGCATGGGTCCGGTTCTG 474
QY 124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTGCTCTTCTCTATCTTCCTC 183
Db 475 GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTGCTCTTCTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGAAGTGTGCC 243
Db 535 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGAAGTGTGCC 594
QY 244 GGGATGTACATGTCAACGAACTGCTCAACTCAAGCATTTGTATGAGGCGGCGAC 303
Db 595 GGGATGTACATGTCAACGAACTGCTCAACTCAAGCATTTGTATGAGGCGGCGAC 654
QY 304 ATGATCATGCAACACCCCGGGTGGTGGCTTCGCGTTCGGGAGAAACAATCTTCCCGCTGC 363
Db 655 ATGATCATGCAACACCCCGGGTGGTGGCTTCGCGTTCGGGAGAAACAATCTTCCCGCTGC 714
QY 364 TGGGTAGCGCTCACCCCGGCGCTGCGAGTGGAGACGCGCTCCCAACAGCAATA 423
Db 715 TGGGTAGCGCTCACCCCGGCGCTGCGAGTGGAGACGCGCTCCCAACAGCAATA 774
QY 424 CGACGCCACGTGCAT----- 438
Db 775 CGACGCCACGTGCATTTGCTGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGG 834
QY 439 -----TCCAGCTTTACCATCTCGCTCGCGGAT 471
Db 835 GACCTCTGCGGATCTGCTTCTCCAGCTGTTCAACATCTCGCTCGCGGAT 894
QY 472 GAGACGGTGCAGACTGCAATTTGCTCAATCTATCCCGGCGCAATACGGGTACCGCTATG 531

Db	895	GAGCGGTGCGAGCTGCAATTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG	954
Qy	532	GCTTGGGATATGATGATGAACCTGGTCGCTCAACAGCGCCCTGGTGTATCGCAGCTGCTC	591
Db	955	GCTTGGGATATGATGATGAACCTGGTCGCTCAACAGCGCCCTGGTGTATCGCAGCTGCTC	1014
Qy	592	CGGATC 597	
Db	1015	CGGATC 1020	
RESULT 14			
Id	AAU48940	standard; DNA; 2434 BP.	
Ac	AAU48940;		
Dt	24-OCT-2002	(first entry)	
XX	Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.		
DE	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;		
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;		
KW	virucide; immunostimulant; vaccine; ds.		
XX	Hepatitis C virus.		
OS	Hepatitis C virus.		
PN	WO20025548-A2.		
PD	18-JUL-2002.		
XX	11-JAN-2002; 2002WO-EP00219.		
PF	11-JAN-2001; 2001US-260699P.		
PR	30-AUG-2001; 2001US-315768P.		
XX	(INNO-) INNOGENETICS NV.		
PA	Maertens G, Bosman F, Buyse M;		
PI	WPI; 2002-599657/64.		
DR	P-PSDB; AAO18679.		
XX	New therapeutic vaccine compositions comprising at least one purified		
PT	recombinant hepatitis C virus (HCV) single or specific oligomeric		
PT	recombinant envelope protein E1 or E2, useful for immunizing humans		
PT	from HCV infection		
XX	Example 2; Page 212-215; 243pp; English.		
PS	The present invention relates to new therapeutic vaccine compositions for		
XX	inducing hepatitis C virus (HCV)-specific antibodies, comprising a		
CC	composition containing at least one purified recombinant HCV single or		
CC	specific oligomeric recombinant envelope proteins selected from an E1 and		
CC	an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are		
CC	useful for inducing HCV-specific antibodies or for immunising humans		
CC	against HCV. The recombinant HCV E1 and/or E2 proteins are useful as		
CC	vaccines or therapeutics, in HCV screening and confirmatory antibody		
CC	tests, for raising antibodies, in the preparation of medicament, and for		
CC	in vitro monitoring of HCV disease or prognosing the response to		
CC	treatment of patients suffering from HCV infection. The present sequence		
CC	is a coding sequence described in the exemplification of the invention.		
XX	Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 other;		
Qy	Query Match	78.8%; Score 501; DB 24; Length 2434;	
XX	Best Local Similarity	89.1%; Pred. No. 1.2e-122;	
XX	Matches 594; Conservative	0; Mismatches 0; Indels 73; Gaps 2;	
Qy	4 TTGGGTAAAGTATCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCCG	63	
Db	355 TTGGGTAAAGTATCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCCG	414	

Mon Dec 22 13:28:43 2003

QY 529 ATGGCTTGGGATATGATGAACTGGT 556
Db 601 ATGGCTTGGGATATGATGAACTGGT 628

Search completed: December 19, 2003, 18:51:21
Job time : 178.828 secs

DR WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di-sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by a
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
XX Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;
XX
Query Match 71.8%; Score 456.4; DB 17; Length 633;
Best Local Similarity 86.8%; Pred. No. 5.3e-111;
Matches 545; Conservative 0; Mismatches 11; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAAGGTCATCGATACCTTACATCGGGCTTCGCCGACCTCGTGGGTTACATT 60
Db 1 ATGTTGGGTAAGGTCATCGATACCTTACATCGGGCTTCGCCGACCTCATGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTT 120
Db 61 CCGCTCGTCGGCGCCCTAGGGGGCTGCCAGAGCCCTGGCGCATGGCGTCCGGTT 120
QY 121 CTGGAGACGGCGTGAATATGCAACAGGGAATTTGCCCGGTTCCTTCTATCTTC 180
Db 121 CTGGAGACGGCGTGAATATGCAACAGGGAATTTGCCCGGTTCCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
Db 181 CTCTTGGCTTTGCTGTCTGTCTGACCAATTCAGCTTCCGCTTATGAAGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGCAGCG 300
Db 241 TCCGGGATGTACCATGTCAAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGCAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGTCCCTGCGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCAACCCCGACGCTCGCAGTAGGAACGGCAGCGTCCCACTACGACA 420
Db 361 TGCTGGGTAGCGCTCAACCCCGACGCTCGCAGTAGGAACGGCAGCGTCCCACTACGACA 420
QY 421 ATACGACGCCACGTGAT----- 438
Db 421 ATACGACGCCACGTGATTTGCTGTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 439 -----TCCAGCTGTTTCAACCATCTCGCTCGCGG 468
Db 481 GGGGATCTCTGCGGATCTGTCTCTCTGCTCCAGCTGTTTCAACCATCTCGCTCGCGG 540
QY 469 CATGAGACGGTGCAGGACTGCAATTCCTATCCCGGCCACATTAACGGGTCAACCGT 528
Db 541 CATGAGACGGTGCAGGACTGCAATTCCTATCCCGGCCACATTAACAGGTCAACCGT 600

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1636.15 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-27
Perfect score: 636
Sequence: 1 ATGTGGTAGGTCATCGA.....ATCACCACTCACTAATAG 636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba.*
2: em_estchum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rpd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	68.2	10.7	488	9 AV755731	AV755731
C 2	54.6	8.6	492	9 AV758366	AV758366
C 3	41.6	6.5	502	12 B1879124	B1879124
C 4	40.6	6.4	275	9 AV835132	AV835132

5	40.6	6.4	402	9 AV392783	AV392783
6	40.6	6.4	551	9 AV392165	AV392165
7	40.6	6.4	552	12 B1996341	B1996341
8	40.6	6.4	584	12 B1727879	B1727879
9	40.6	6.4	1201	13 BX356664	BX356664
10	40.2	6.3	1162	12 BM918259	BM918259
11	40.2	6.3	1201	13 BX356664	BX356664
C 12	40	6.3	1201	9 AL513886	AL513886
C 13	39	6.1	359	12 BJ252669	BJ252669
C 14	39	6.1	375	12 BJ246716	BJ246716
15	39	6.1	840	29 CC335916	CC335916
16	39	6.1	873	14 CD446071	CD446071
17	38.6	6.1	925	29 CNS0091P	CNS0091P
C 18	38.4	6.0	606	9 AV915997	AV915997
C 19	38.4	6.0	636	12 B1960110	B1960110
20	38.4	6.0	702	14 CD432549	CD432549
21	38.4	6.0	746	9 AV921112	AV921112
22	38.4	6.0	970	29 CNS010C9	CNS010C9
C 23	38.4	6.0	987	29 CNS015VX	CNS015VX
C 24	38.2	6.0	533	6 AU192776	AU192776
C 25	38.2	6.0	538	6 AU193705	AU193705
C 26	38.2	6.0	544	6 AU190971	AU190971
C 27	38.2	6.0	544	6 AU192419	AU192419
C 28	38.2	6.0	1270	12 BG968359	BG968359
C 29	38	6.0	354	14 CB968525	CB968525
30	38	6.0	1201	13 BX381961	BX381961
31	37.8	5.9	435	14 C72860	C72860
32	37.8	5.9	533	29 CC010084	CC010084
33	37.8	5.9	659	29 CC405164	CC405164
34	37.8	5.9	826	29 B2736582	B2736582
C 35	37.8	5.9	895	29 CC359028	CC359028
C 36	37.8	5.9	925	29 CC359026	CC359026
C 37	37.8	5.9	940	29 CC010085	CC010085
C 38	37.8	5.9	951	29 CC405167	CC405167
39	37.6	5.9	431	9 AV639153	AV639153
C 40	37.4	5.9	360	9 AJ473805	AJ473805
C 41	37.4	5.9	637	13 BQ293470	BQ293470
C 42	37.4	5.9	641	13 BQ172543	BQ172543
C 43	37.4	5.9	650	14 CA828039	CA828039
44	37.4	5.9	834	29 B2641450	B2641450
C 45	37.4	5.9	841	29 B2641457	B2641457

ALIGNMENTS

RESULT 1
AV755731/c 488 bp mRNA linear EST 19-OCT-2000
LOCUS AV755731 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES


```

source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAK93"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
116 a 134 c 137 g 97 t 4 others

BASE COUNT
ORIGIN

Query Match 10.7%; Score 68.2; DB 9; Length 488;
Best Local Similarity 72.3%; Pred. No. 4.8e-07;
Matches 102; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 445 CTGTTCCACCATCGCTCGCGGCATGAGACGGTGCAGAGCTGCAATGCTCAATCAT 504
DB 403 CAGCTGATCATCTGGCTTCAGCACCATGAGTTTGTCATGAATGCAATGCTCCATCAT 344
QY 505 CCCGGCCACATACCGGTCACCGTATG-GCTTGGGATATGATGATGAACCTGTCGCCCTAC 563
DB 343 CCTGGGCCATCATCTGCACACCGTATGACATGGGACATGATGATGAAGAACTGTCGTGCAC 284
QY 564 AACGGCCCTGGTGGTATCGCA 584
DB 283 CGCTGCTATGATCATGGCGTA 263

RESULT 2
AV758366/c
LOCUS AV758366 BM Homo sapiens CDNA clone BMFAKA03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens CDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. .492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAKA03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
124 a 128 c 125 g 112 t 3 others

BASE COUNT
ORIGIN

Query Match 8.8%; Score 54.6; DB 9; Length 492;
Best Local Similarity 67.2%; Pred. No. 0.0015;
Matches 92; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

```


Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

BI996341
EST. GI:16431115
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 552)
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermodt, J. P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers
1. 552
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II"
/note="vector: pBluescript II SK; Site_1: EcoRI; Site_2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hr) H2 production conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaudo et al., (1996) Genome Research 6: 791-806."
Research 6: 791-806."

BASE COUNT 93 a 184 c 189 g 86 t
ORIGIN
Query Match 6.4%; Score 40.6; DB 12; Length 552;
Best Local Similarity 45.3%; Pred. No. 6.2; Indels 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
QY 42 CGACCTCGTGGGGTACATTCCGCTCGTGGCGCCGCCCTTAGGGGGCGTGCAGGGCCCT 101
DB 110 CGAGCTCATCTCGTGGTTCGCGGGCACTGCCAACATGAAGGACGTGCTGACGACCT 169
QY 102 GCGCATGCGCTCCGGGTTCTGGAGACGCGGTGAATATGCAACAGGGAATTCGCCGG 161
DB 170 GCGCGCGCGCGCGAGTGGAGGGCGGCTACGCGCACGAGTCGGTAGCTTGGGCGC 229
QY 162 TTGCTCTTTCTATCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 221
DB 230 CCACAAGGTGTTGACGAGATCAAGAGTACGTGTGAACCTCAGGCCACAGACCCAG 289
QY 222 TTATGAAGTGCACACGCTGTCGGGATGTACCATGTACGAACGATGCTCTCCACTCAAG 281
DB 290 CTTGCGCGTCCGCTGCGTGGCCACTCGCTGGCGCGGCAACCGCGGCTGCTGTGAT 349
QY 282 CATGTGTATGAGCGACGACATGATCATGACACCCCGGGTGGTGGCTTCCGTTCCG 341
DB 350 CCGTATGACACACGACGAGAGTTTGGCGCGCATCTACGCGCGGTGTCCTCATGCGGG 409

AV392165 551 bp mRNA linear EST 23-APR-2002
AV392165 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CM083605_r 5', mRNA sequence.
AV392165
EST. GI:6546381
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 551)
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags
DNA Res. 6 (6), 369-373 (1999)
20152988
10691129
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1. 551
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM083605_r"
/dev_stage="Photoautotrophic growth"
/clone_lib="Chlamydomonas reinhardtii C9"
/note="vector: pBluescriptII SK; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 94 a 182 c 189 g 85 t 1 others
ORIGIN
Query Match 6.4%; Score 40.6; DB 9; Length 551;
Best Local Similarity 45.3%; Pred. No. 6.2; Indels 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
QY 42 CGACCTCGTGGGGTACATTCCGCTCGTGGCGCCGCCCTTAGGGGGCGTGCAGGGCCCT 101
DB 108 CGAGCTCATCTCGTGGTTCATGTGCGCGCACTGCCAACATGAAGGACGTGCTGACGACCT 167
QY 102 GCGCATGCGCTCCGGTTCGAGGACGCGGTGAACATATGCAACAGGGAATTTGCCCGG 161
DB 168 GCGCGCGCGCGCGAGTGGAGGGCGGTACGCGCAGAGTCCGTGAGCTTGGGCGC 227
QY 162 TTGCTCTTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 221
DB 228 CCGAAGGTGTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCACAGAACCCAG 287
QY 222 TTATGAAGTGCACACGCTGTCGGGATGTACCATGTACGACGACCTGCCAATCAAG 281
DB 288 CTTGCGCGTCCGCTGCGTGGGCCACTCGCTGGGCGGGGCAACCGCGGCTGCTGTGAT 347
QY 282 CATGTGTATGAGCGACGACATGATCATGACACCCCGGGTGGTGGCTTCCGTTCCG 341
DB 348 CCGTATGACACACGACGAGAGTTTGGCGCGGATCTACGCGCGGTGTCCTCATGCGGG 407
QY 342 GCGAAGCAACTCTTCCCGCTGCTGGGT 368
DB 408 CAAGAAGACGACGAGGAGTACATGAT 434

RESULT 7
BI996341 552 bp mRNA linear EST 25-OCT-2001
LOCUS 1031037A07.y2 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION


```

QY 342 GGAGACAACTCTCCCGCTGCTGGGT 368
Db |||||
410 CAAGAAGACGAGGCGAGCTACATGAT 436

RESULT 8
B1727879
LOCUS B1727879 584 bp mRNA linear EST 19-SEP-2001
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION B1727879
VERSION B1727879.1 GI:15703574
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 584)
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
JOURNAL Unpublished
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
source
1..584
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoRI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
BASE COUNT 106 a 188 c 197 g 93 t
ORIGIN
Query Match 6.4%; Score 40.6; DB 12; Length 584;
Best Local Similarity 45.3%; Pred. No. 6.3;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGTACATTCGCTCGCGCGCCCTAGGGGGCGTCCAGGCGCCT 101
Db |||||
47 CGAGCTCATCTGTCATTTGTCGCGGCACTGCCAACATGAGGAGCTGCTGACGACCT 106

QY 102 GCGCATGCGCTCGGGTTCTGAGGACGCGGTGAATGCAACAGGGAATTTGCCCGG 161
Db |||||
107 GCGCGCGCGCGCGAGTGGAGGCGGTACGCGCACGAGTCCGTGAGCTTGGCGC 166

QY 162 TTGCTCTTCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 221
Db |||||
167 CCGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCCCAGACCCAG 226

```

```

QY 222 TTATGAAGTGCACAACTGTTCCGGGATGACCATGTACGAACGACTGCTCCAACCTCAAG 281
Db |||||
227 CTTGCGCGTCCGCTGCTGGGCACTCGCTGGCGGCGCACCGCGGCTGCTGCTGCTGAT 286

QY 282 CATTTGTTATGAGGAGGAGGACATCATGTCACACCCCGGGTGGTGGTGGTGGTGG 341
Db |||||
287 CTTGATGACCAACGAGGAGGAGTTCGCGCGGCGCATCTACGCGGCGGTGCTCCATGCGCGG 346

QY 342 CGAGAACAACTCTCCCGCTGCTGGGT 368
Db |||||
347 CAAGAAGACGAGGCGAGCTACATGAT 373

RESULT 9
B1727879
LOCUS B1727879 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION Clone CSODI015YB03 3-PRIME, mRNA sequence.
ACCESSION B1727879
VERSION B1727879.1 GI:30378083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI015CA02NP1.

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI015YB03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN
Query Match 6.4%; Score 40.4; DB 13; Length 1201;
Best Local Similarity 10.4%; Pred. No. 8.8;
Matches 52; Conservative 233; Mismatches 212; Indels 3; Gaps 1;

QY 34 GCGTTCGCGACCTCGTGGGTACATTCGCTCGCGCGCCCTAGGGGGCGCTGCC 93
Db |||||
618 GNTNTSSSSSTNNNNSSSSNNNTNTTBTBTSSSTSSSTSSSTSSSTSSSTSSST 677

QY 94 AGGCGCCTCGCGCATGCGCTCGGGTCTGAGGAGCGCTGAACATATCAACAGGGAAT 153
Db |||||
678 SSSSSSBTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 737

QY 154 TTGCCCGGTGCTCTTCTATCTTCCTCTTGGCTTGTCTGCTGCTGCTGCTGCTGCT 213
Db |||||
738 TTKSSSSSTBSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 797

QY 214 GCTTCGCTGATGAGTGGCAACGTGTCGCGGATGTACCATGTACGAACACTGCTCC 273
Db |||||
798 TBSMTSSSBTCTSSSSSSSBTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 954

```


[illegible]

RESULT 10		
BM918259	linear	EST 12-MAR-2002
LOCUS	1162 bp	mRNA
DEFINITION	AGENCOURT_6611605 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485649	
	5', mRNA sequence.	
ACCESSION	BM918259	GI:19368638
VERSION	BM918259.1	GI:19368638
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 1162)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished	
COMMENT	Contact: Robert Straubeberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Dr. Daniel McVicar, DBS/NCI cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2016 row: n column: 18 High quality sequence stop: 567.	

```

FEATURES
source
Location/Qualifiers
1. .1162
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5485649"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
224 a 499 c 240 g 198 t 1 others
BASE COUNT
ORIGIN
Query Match 6.3%; Score 40.2; DB 12; Length 1162;
Best Local Similarity 54.4%; Pred. No. 9.8;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

Db	715	CCCCCCCCGGTCCCTGCCCCACACCCCGGACCCCAAAACCCCCCGGACCGCTCC	774
Qy	373	CTCACCCCGCGCTCGAGCTAGGAAGCCGAGGTCGCCACACGACAATACGACGCCAC	432
Db	775	CCCACTGCCACCGCACCCGCCCATATCGCCCGCTACCCGATCAGCACTACCCCCAC	834
Qy	433	GTGCGATCCCCAGCTGTTCACCATCTCGCC	461
Db	835	GCTCGATCCCGGCGCTGCACACCCCGCC	863

RESULT 11
BX356664/C

LOCUS	1201 bp	linear	EST 05-MAY-2000
DEFINITION	25-NORMALIZED Homo sapiens CDNA		
DEFINITION	3-PRIME mRNA sequence.		
LOCUS	1201 bp	linear	EST 05-MAY-2000
DEFINITION	25-NORMALIZED Homo sapiens CDNA		
DEFINITION	3-PRIME mRNA sequence.		

Accession	BX356664	clone CS0DIVISIB03_3-FR
Insertion	PY356664_1	GI:30378083

VERSION BYJ00001.1
KEYWORDS EST.

SOURCE	Homo sapiens (human)
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE: 1 (bases 1 to 1201)

AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE	Full-length cDNA libraries and normalization
ACCESSION	U98671
KEYWORDS	transcriptomics; gene expression; differential expression; cDNA library; sequencing; RT-PCR
ABSTRACT	The purpose of this study was to generate full-length cDNA libraries from total RNA samples and to normalize them for use in differential expression analysis. The libraries were constructed using a modified method of cDNA synthesis and ligation. The libraries were normalized by size exclusion chromatography and then used for differential expression analysis. The results showed that the normalized libraries were more effective than the non-normalized libraries in identifying differentially expressed genes.
INTRODUCTION	Differential expression analysis is a powerful tool for studying gene expression patterns. It allows researchers to identify genes that are up-regulated or down-regulated under specific conditions. One of the most common methods for differential expression analysis is the use of cDNA libraries. However, traditional cDNA libraries often contain many non-expressed or lowly expressed genes, which can obscure the signal of interest. To overcome this problem, several methods have been developed to improve the quality of cDNA libraries. One such method is the use of full-length cDNA libraries, which contain only full-length transcripts. Another method is the use of normalized cDNA libraries, which have been adjusted to represent equal amounts of each transcript. In this study, we compared the effectiveness of full-length cDNA libraries and normalized cDNA libraries in differential expression analysis.
MATERIALS AND METHODS	Total RNA was isolated from cells treated with various agents. The RNA was then reverse-transcribed into cDNA using a modified protocol. The cDNA was ligated into a vector and transformed into bacteria. The resulting cDNA libraries were normalized by size exclusion chromatography. The libraries were then used for differential expression analysis using RT-PCR.
RESULTS	The results of the differential expression analysis showed that the normalized libraries were more effective than the non-normalized libraries in identifying differentially expressed genes. Specifically, the normalized libraries identified a greater number of up-regulated and down-regulated genes.
CONCLUSION	This study demonstrates that normalized cDNA libraries are more effective than non-normalized libraries in differential expression analysis. This finding has important implications for the design of cDNA libraries for gene expression studies.
REFERENCES	1. Schena M, Alizadeh AA, Sheng L, et al. (2000) Large-scale differential display of complementary DNAs using microarrays. <i>Nature Genet</i> 20:289-296. 2. Schena M, Heller J, Ching MA, et al. (1995) Parallel processing of DNA arrays. <i>Science</i> 268:475-478. 3. Schena M, Heller J, Ching MA, et al. (1996) A microarray-based approach to genome-wide differential screening. <i>Proc Natl Acad Sci USA</i> 93:10614-10619. 4. Schena M, Heller J, Ching MA, et al. (1997) A microarray-based approach to genome-wide differential screening. <i>Proc Natl Acad Sci USA</i> 94:10614-10619. 5. Schena M, Heller J, Ching MA, et al. (1998) A microarray-based approach to genome-wide differential screening. <i>Proc Natl Acad Sci USA</i> 95:10614-10619.

JOURNAL
Unpublished
Contact: Genoscope
COMMENT

Genoscope - Centre National de Séquençage

BP 191 91006 EVRY cedex - France
 Email: cecile@genoscope.cns.fr Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : www.invitrogen.com / Invitrogen Corporation 1600

http://fulllength.invitrogen.com/
 Transduction: Genoscope sequence ID : CS0DI015CA02NP1.

FEATURES

FEATURED
source
1. .1201

```
/organism="Homo sapiens"
/mol type="tRNA"
```

```
/db xref="taxon:9606"
```

```

/clone="CS0DI015YB03"
COT 25-NORMALIZED"

```

```
/tissue_type="PLACENTA COT 25-NORMALIZED"  
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
```

```

/clone=115- name=115-
/note="1st strand cDNA was primed with a NotI-oligo(dT)

```

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT	116 a	88 c	93 g	398 t	506 others
BASE COUNT	116 a	88 c	93 g	398 t	506 others

ORIGIN

Query Match	6.3%	Score 40.2;	DB 13;	Length 1201;
Best Local Similarity	11.4%	Pred. No. 9.9;	Indels 0;	Gaps 0;
Matches 46;	Conservative 186;	Mismatches 171;		

QY
227 AAGTGGCAACGTGTCGGGATGTACCATGTACGAACACTGTCCAACTCAACATTG 286

||| : : : : ||| : : : : ||| : : : : ||| : : : :
|||| : : : : ||| : : : : ||| : : : : ||| : : : : ||| : : : : ||| : : : :

D6	1185	AASSVAASSSSSSSSVSAAAAAASATVAAAAAASASSSSSSVVVSAAASAAAACCCGCGAG
Ov	287	TGTATGAGGCAGCGACAATGATCATGCAACCCCCGGTGCCTGCCCTTCGGTTCCGGAGA 346

[illegible][illegible]

QY 407 TCCTCCACGACGAATACGACGCCACGTGGATTCCGAGCTGTTACCATCTCGCTCGCC 466

Db	1005	AAAASSVAAAVSRAASVSABAAAAAASVAAASSSSSSSVSSSSSSSAVSSSSSSSSSVSSSSSS	346
Ov	467	GGCATGACAGCGTGACGACTGCAATTCGTCAATCTATCCGGCCACATACGGGTCAACC	526

[illegible]

Db 945 AVSASSSSSSAAASVSSAAVSSSSVAAASSSSSSSSVAAASSSSSSAAASVAVSAAAV 886

QY 527 GTATGCTGGGATATGATGAACTGGTGCCTTACACGCCCTGGTGGTATCGCAGC 586

Db 885 VSWSSVASSASVASVSSSSSSAAASVSSSSAAASVSSSSAAASVSSSSAAASVSSMAS 826

QY 587 TGCTCGGATCGTATCGAGGCGACACCATCACCACCATCA 629

Db 825 AAASSAAVSSSSSSSAGAVSSSAKSVASSASVSSAGSSSA 783

RESULT 12

AL513886/LOCUS

AL513886 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA0062G08

DEFINITION 5-PRIME, mRNA sequence.

ACCESSION AL513886

VERSION AL513886.2 GI:30463771

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 13, 2001 this sequence version replaced gi:12777380.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4924.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/Cluster.cgi?seq=CL0BA0062G08P1&cluster=4924.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CL0BA0062G08P1.

FEATURES

source

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CL0BA0062G08"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 201 a 311 c 349 g 146 t 194 others

ORIGIN

Query Match 6.3%; Score 40; DB 9; Length 1201;

Best Local Similarity 26.5%; Pred. No. 11;

Matches 103; Conservative 104; Mismatches 179; Indels 3; Gaps 1;

QY 32 GCGGCTTCCGACCTCGTGGGATACATTCGCTCGTCGCGCCCCCTAGGGGGCGT 91

Db 1129 GSGCSCGTGGCSBSCTYKKGKBSSSCCSVSSSGSGSCSSSCSCCGGGSGGGGG 1070

QY 92 CCAGGCGCCCTGGCGCATGCGGTCTCGGATCGGAGACGCGTGAACACAGGA 151

Db 1069 CCSCGKGGGGKSSSGSGGCCCGCGG---GGSSSGSGSGGSCCWAASSSAYKKKK 1013

QY 152 ATTGCGCGGTGCTCTCTATCTCTATCTCTCTCTGCTTGTGCTCTGACCGTTC 211

Db 1012 GKKTCTTTTMMAAATHTTTTWTTTTCTTAAGGGGTAKVAKCCWCCCGCCGAGTTS 953

QY 212 CAGCTTCCGCTTATGAAGTGGCAACGCTGTCGGGATGTACCATGTCAAGAACGACTGCT 271

Db 952 GACSCSCGCAADAACVCGAGMDSGAMKGTGVGSCCTTSRCKWGGGTTSGMMVGCATTY 893

QY 272 CCAACTCAAGCATTTGTATGAGGCGGAGCATCATGACACACCCCGCGGTGGTGC 331

Db 892 AYBSYTGTRRTWTGTTGCTYASGSGMWSSKRBKKGCCMAYAACSCGAGASCST 833

QY 332 CTTGCGTTGGGAGAACAACTTTCCCGTGTGGGTAGCGTCAACCCACGCTCGCAG 391

Db 832 SGCSKGTTKTTTGTGCTGTAAGASMAAGRTWAGGGGGGGGGCCCGCCCGCCCVB 773

QY 392 CTAGGAAGCCGAGCGTCCCGACACGACA 420

Db 772 BBCCCMCHCTKSCRCMCCRGACTYCCCCA 744

RESULT 13

BJ252669/LOCUS

BJ252669 Y. Ogiwara unpublished cDNA library, Wh_f Triticum

DEFINITION aestivum cDNA clone whf25g19 3', mRNA sequence.

ACCESSION BJ252669

VERSION BJ252669.1 GI:20061830

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

AUTHORS Ogiwara, Y. and Murai, K.

TITLE Expressed genes in Triticum aestivum

JOURNAL Unpublished

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..359

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="whf25g19"

/tissue_type="spike at flowering date"

/dev_stage="Feekes' scale 10.5.1"

/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"

BASE COUNT 70 a 115 c 107 g 67 t

ORIGIN

Query Match 6.1%; Score 39; DB 12; Length 359;

Best Local Similarity 58.0%; Pred. No. 14;

Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCCGCTCGCGGCCCGCCCTAGS 83

Db 297 CTTCAAGTCAACAGCGCGCTCTGGAAGCGCTCAGGGCGGTGCGCGCGTCCGCGG 238

QY 84 GGGCGCTGCGAGGGCCCTCGCGCATGCGCTCCGGGTTCTGGAGGAGCGGCTGAACCTATG 142

Db 237 GGACGCCGACGCCCTTGGCGCAGACGCTGACGCTGCTGCCGTTGACGTCGCCCAAG 179

RESULT 14

BJ246716/LOCUS

BJ246716 Y. Ogiwara unpublished cDNA library, Wh_f Triticum

DEFINITION aestivum cDNA clone whf25g19 5', mRNA sequence.

ACCESSION BJ246716

VERSION BJ246716.1 GI:20058228

Mon Dec 22 13:28:45 2003

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

REFERENCE
1 (bases 1 to 375)
Ogihara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..375
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf5g19"
/tissue_type="spike at flowering date"
/dev_stage="Peekes", scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cdna library, Wh_f"
BASE COUNT      81 a   107 c   110 g   77 t
ORIGIN
Query Match      6.1%; Score 39; DB 12; Length 375;
Best Local Similarity 58.0%; Pred. No. 14;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 24 CCTACATCGCGTTCCGACCTCGTGGGTACATTCCGCTCGTGGGGCCCCCTAGG 83
Db 36 CTTCAAGTGAACAGCCCGTCTGGAAGCGCTCAGGGCGGTGACGCGCGTGG 95

Qy 84 GGGCGCTGCCAGGCGCCCTGGCGCATGCGTCCGGTTCTGGAGGAGCGGCGTGAATG 142
Db 96 GGACCGCGGAGCCCTCGGGCGCAGGACGTGACGCTGCTGCGGTCGACGTCGCGCAAGG 154

RESULT 15
LOCUS      CG335916      840 bp      DNA      linear      GSS 16-MAY-2003
DEFINITION      OGUAJ60TV ZM.0.7.1.5 KB Zea mays genomic clone ZMMEMa0393124,
                  genomic survey sequence.
ACCESSION      CC335916
VERSION      CC335916.1 GI:30805329
KEYWORDS      GSS.
SOURCE      Zea mays
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick
, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek
, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..840
/organism="Zea mays"

FEATURES
source

```

```

/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMEMa0393124"
/clone_lib="ZM.0.7.1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
BASE COUNT      103 a   298 c   291 g   148 t
ORIGIN
Query Match      6.1%; Score 39; DB 29; Length 840;
Best Local Similarity 53.6%; Pred. No. 18;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 39 CGCGACCTCGTGGGTACATTCCGCTCGTGGGGCCCCCTAGGGGGGCTGCCAGGGC 98
Db 292 CCCGGCGTCGACGGCGCCACTCCGTTCACTGCGCGGCTCGCGGGAACGC 351

Qy 99 CTTGGCGCATGGGCTCCGGTTCTGGAGGACGCGTGAATGCAACAGGGGAATTGGC 158
Db 352 TGTGGCGTCTCAAGCTGCTCTGGGTGCGGGCGGACCCGGTAACCTCCGACTCCGC 411

Qy 159 CGGTGCTCTTTCTCTATCTTCTCTTGGCT 189
Db 412 CGGTGCTTGGCTGCGGACGTCATCTTGTCT 442

Search completed: December 20, 2003, 06:54:51
Job time : 1638.15 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 45.8681 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-27
Perfect score: 636
Sequence: 1 ATGTTGGTAGGTCATCGA.....ATCACCACCACTACTAATAG 636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	636	100.0	636	3	US-08-612-973-27
2	636	100.0	636	3	US-08-927-597-27
3	597.4	93.9	606	3	US-08-612-973-25
4	597.4	93.9	606	3	US-08-927-597-25
5	597	93.9	723	3	US-08-612-973-21
6	597	93.9	723	3	US-08-927-597-21
7	556	87.4	561	3	US-08-612-973-23
8	556	87.4	561	3	US-08-927-597-23
9	515	81.0	795	3	US-08-612-973-5
10	515	81.0	795	3	US-08-927-597-5
11	512	80.5	2082	3	US-08-612-973-47
12	512	80.5	2082	3	US-08-927-597-47
13	512	80.5	2433	3	US-08-612-973-49
14	512	80.5	2433	3	US-08-927-597-49
15	456.4	71.8	636	3	US-08-612-973-7
16	456.4	71.8	636	3	US-08-927-597-7
17	451.6	71.0	636	3	US-08-612-973-13
18	451.6	71.0	636	3	US-08-927-597-13
19	448	70.4	1539	2	US-08-470-4268-17
20	448	70.4	1863	2	US-08-470-4268-14
21	443.2	69.7	9595	3	US-09-014-416-4
22	443.2	69.7	9599	3	US-09-014-416-6
23	441.6	69.4	932	1	US-08-081-072-15
24	441.6	69.4	932	1	US-08-449-093A-15
25	441.6	69.4	2116	3	US-08-191-160-21
26	440	69.2	1037	1	US-08-462-195-1
27	440	69.2	1037	2	US-08-636-883-1

28	440	69.2	1037	3	US-09-127-829-1	Sequence 1, Appli
29	440	69.2	9472	4	US-08-150-204E-96	Sequence 96, Appli
30	428.8	67.4	1167	1	US-08-324-977-9	Sequence 9, Appli
31	428.8	67.4	1167	2	US-08-384-616-9	Sequence 9, Appli
32	428.8	67.4	1167	2	US-08-904-686A-9	Sequence 9, Appli
33	428.8	67.4	1167	3	US-09-315-850-9	Sequence 9, Appli
34	428.8	67.4	1499	1	US-08-324-977-3	Sequence 3, Appli
35	428.8	67.4	1499	2	US-08-384-616-3	Sequence 3, Appli
36	428.8	67.4	1499	2	US-08-904-686A-3	Sequence 3, Appli
37	428.8	67.4	1499	3	US-09-315-850-3	Sequence 3, Appli
38	428.8	67.4	6039	1	US-08-324-977-11	Sequence 11, Appli
39	428.8	67.4	6039	2	US-08-384-616-11	Sequence 11, Appli
40	428.8	67.4	6039	2	US-08-904-686A-11	Sequence 11, Appli
41	428.8	67.4	6039	3	US-09-315-850-11	Sequence 11, Appli
42	428.8	67.4	9030	1	US-08-324-977-13	Sequence 13, Appli
43	428.8	67.4	9030	2	US-08-384-616-13	Sequence 13, Appli
44	428.8	67.4	9030	2	US-08-904-686A-13	Sequence 13, Appli
45	428.8	67.4	9030	3	US-09-315-850-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-08-612-973-27
; Sequence 27, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLSCULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630

US-08-612-973-27	Query Match	100.0%;	Score 636;	DB 3;	Length 636;	
	Best Local Similarity	100.0%;	Pred. No. 1.7e-157;			
	Matches 636;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGTTGGGTAAGGTCATCGATACCTTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60			
Db	1	ATGTTGGGTAAGGTCATCGATACCTTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60			
Qy	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTCCAGGGCCCTTGGCGCATGGCGTCCGGGTT	120			
Db	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTCCAGGGCCCTTGGCGCATGGCGTCCGGGTT	120			
Qy	121	CTGGAGGACGGCGTGAACTATGCAACAGGGAAATTGCCCGGTGCTCTTCTCTATCTTC	180			
Db	121	CTGGAGGACGGCGTGAACTATGCAACAGGGAAATTGCCCGGTGCTCTTCTCTATCTTC	180			
Qy	181	CTCTTGGCTTTTCTGTCGTCTGACCGTTCCAGCTTCGGCTTATGAAGTCGCAACGTG	240			
Db	181	CTCTTGGCTTTTCTGTCGTCTGACCGTTCCAGCTTCGGCTTATGAAGTCGCAACGTG	240			
Qy	241	TCCGGGATGTACCATGTGTCAAGAACGACTGTCCAACTCAAGCATTTGTATGAGCAGCG	300			
Db	241	TCCGGGATGTACCATGTGTCAAGAACGACTGTCCAACTCAAGCATTTGTATGAGCAGCG	300			
Qy	301	GACATGATCATGCACACCCCGGGTGGCGCCCTGCGTTCCGGGAGAAACAATCTTCCCGC	360			
Db	301	GACATGATCATGCACACCCCGGGTGGCGCCCTGCGTTCCGGGAGAAACAATCTTCCCGC	360			
Qy	361	TGCTGGGTAGCGCTCACCCGCCACGCTCGCAGCTTAGGAACCCAGCGTCCCAACCAAGCA	420			
Db	361	TGCTGGGTAGCGCTCACCCGCCACGCTCGCAGCTTAGGAACCCAGCGTCCCAACCAAGCA	420			
Qy	421	ATACGACGCCACGTCGATTCCGAGCTGTTCCATCTCGCCTTCGCGGATGAGACGGTG	480			
Db	421	ATACGACGCCACGTCGATTCCGAGCTGTTCCATCTCGCCTTCGCGGATGAGACGGTG	480			
Qy	481	CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACCGGTCAACCGTATGGCTTGGGAT	540			
Db	481	CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACCGGTCAACCGTATGGCTTGGGAT	540			
Qy	541	ATCATGATGAACGTGTCGCCCTACACGGCCCTGGTGATTCGAGCTGCTCCGATCGTG	600			
Db	541	ATCATGATGAACGTGTCGCCCTACACGGCCCTGGTGATTCGAGCTGCTCCGATCGTG	600			
Qy	601	ATCAGGGCAGACCACTATCACCACCATCATATAG	636			
Db	601	ATCAGGGCAGACCACTATCACCACCATCATATAG	636			

RESULT 2
US-08-927-597-27
; Sequence 27, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

, OPERATING SYSTEM: PC-DOS/MS-DOS
,
, SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
,
, CURRENT APPLICATION DATA:
,
, APPLICATION DATA:
,
, APPLICATION NUMBER: US/08/927,597
,
, FILING DATE:
,
, CLASSIFICATION:
,
, PRIOR APPLICATION DATA:
,
, APPLICATION NUMBER: US 08/612,973
,
, FILING DATE: 11-MAR-1996
,
, ATTORNEY/AGENT INFORMATION:
,
, NAME: BYRNE, THOMAS E.
,
, REGISTRATION NUMBER: 32,205
,
, REFERENCE/DOCKET NUMBER: 1497-10
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: (703) 816-4000
,
, TELEFAX: (703) 816-4100
,
, INFORMATION FOR SEQ ID NO: 27:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 636 base pairs
,
, TYPE: nucleic acid
,
, STRANDEDNESS: single
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: cdna
,
, HYPOTHETICAL: NO
,
, ANTI-SENSE: NO
,
, FEATURE:
,
, NAME/KEY: CDS
,
, LOCATION: 1..633
,
, FEATURE:
,
, NAME/KEY: mat peptide
,
, LOCATION: 1..630
,
, US-08-927-597-27

```

Query Match	100.0%;	Score 636;	DB 3;	Length 636;
Best Local Similarity	100.0%;	Prod. No. 1.7e-157;	Indels	Gaps
Matches 636;	Conservative	0;	Mismatches	0;
QY	1	ATGTTGGGTAAAGTCATCGATACACCTTTACATGCGGCTTCGCCGACCTCGCTGGGGGTACATT	60	
DB	1	ATGTTGGGTAAAGTCATCGATACACCTTTACATGCGGCTTCGCCGACCTCGCTGGGGGTACATT	60	
QY	61	CCGCTCGTCCGCGCCCCCTTAGGGGGCGCTGCGCAGGGCCCTGGCGCATGGCGTCCGGGTT	120	
DB	61	CCGCTCGTCCGCGCCCCCTTAGGGGGCGCTGCGCAGGGCCCTGGCGCATGGCGTCCGGGTT	120	
QY	121	CTGAGAGCAGCGGTGAACATGACACAGGGAATTTGCCCGGTTCGCTCTTTCTCTATCTTC	180	
DB	121	CTGAGAGCAGCGGTGAACATGACACAGGGAATTTGCCCGGTTCGCTCTTTCTCTATCTTC	180	
QY	181	CTCTTGGCTTTTGCTGTCCTGTCCTGACCCGTTCCAGCTTCGCGTTATGAAGTGGCAACGTG	240	
DB	181	CTCTTGGCTTTTGCTGTCCTGTCCTGACCGTTCCAGCTTCGCGTTATGAAGTGGCAACGTG	240	
QY	241	TCCGGGATGTACCATGTCACGAAAGACTGCTCCAACCTCAAGCATTTGTGTAGGCGAGCG	300	
DB	241	TCCGGGATGTACCATGTCACGAAAGACTGCTCCAACCTCAAGCATTTGTGTAGGCGAGCG	300	
QY	301	GACATGATCATGCAACACCCCGGGTCGTCCTCGGTTCCGGAGAAACAACCTCTTCCCGC	360	
DB	301	GACATGATCATGCAACACCCCGGGTCGTCCTCGGTTCCGGAGAAACAACCTCTTCCCGC	360	
QY	361	TGCTGGGTAGCGCTCAACCCCGACGCTCGCAGTATGGAAAGCCAGCGTCCCAACACGACA	420	
DB	361	TGCTGGGTAGCGCTCAACCCCGACGCTCGCAGTATGGAAAGCCAGCGTCCCAACACGACA	420	
QY	421	ATACGACGCCAGCTCGATTCACGCTGTTTCAACATCTCGCCCTCGCCGGCATGAGACGGTG	480	
DB	421	ATACGACGCCAGCTCGATTCACGCTGTTTCAACATCTCGCCCTCGCCGGCATGAGACGGTG	480	
QY	481	CAGGACTGCATTTGCTCAATCTATCCGGCCACATAACGGGTTCACCGTATGGCTTGGGAT	540	
DB	481	CAGGACTGCATTTGCTCAATCTATCCGGCCACATAACGGGTTCACCGTATGGCTTGGGAT	540	

QY	1	ATGTTGGGTAAAGTTCATGATACCCCTTACATCGGCTTCGGCGA	CCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAAGTTCATGATACCCCTTACATCGGCTTCGGCGA	CCTCGTGGGGTACATT	60
QY	61	CGGTCGTGCGGCGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATGGCGTC	CGGGTT	120
Db	61	CGGTCGTGCGGCGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATGGCGTC	CGGGTT	120
QY	121	CTGGAGGACGGCGTGAACTATGCAACAGGGAAATTTGCCGGTTGCTCTTCTCTATCTTC		180
Db	121	CTGGAGGACGGCGTGAACTATGCAACAGGGAAATTTGCCGGTTGCTCTTCTCTATCTTC		180
QY	181	CTCTTGGCTTTGCTGTCCTGTCTGA	CGGTTCCAGCTTCGGCTATGAAGTGGCGCAACGTG	240
Db	181	CTCTTGGCTTTGCTGTCCTGTCTGA	CGGTTCCAGCTTCGGCTATGAAGTGGCGCAACGTG	240
QY	241	TCCGGATGTACCATGTCA	CGAACGACTGCTCCAACTCAAGCAATTTGTGATGAGGCAGCG	300
Db	241	TCCGGATGTACCATGTCA	CGAACGACTGCTCCAACTCAAGCAATTTGTGATGAGGCAGCG	300
QY	301	GACATGATCATGCACACCCCGGGTGGTGCCCTTGGGAGAACAACTCTTCCCGC		360
Db	301	GACATGATCATGCACACCCCGGGTGGTGCCCTTGGGAGAACAACTCTTCCCGC		360
QY	361	TGCTGGGTAGCGCTCACCCCCACGCTCGCAGT	AGGAAACGCCAGCGTCCCAACCAACGACA	420
Db	361	TGCTGGGTAGCGCTCACCCCCACGCTCGCAGT	AGGAAACGCCAGCGTCCCAACCAACGACA	420
QY	421	ATACGAGGCCACGTCGATTC	CCAGGTGTTACCAATCTCGCCTCGCGGCATGAGACGGTGG	480

Db 421 ATACGAGCCAGTCGATTCCAGCTGTTTCCACATCTCCCTCGCGGGCATGAGACGGTG 480
QY 481 CAGGAGTCAATGTCTCAATCTATCCCGCCACATAACGGGTACCGTATGCTTGGGAT 540
Db 481 CAGGAGTCAATGTCTCAATCTATCCCGCCACATAACGGGTACCGTATGCTTGGGAT 540
QY 541 ATGATGATGAATGCTGCGCTACAGCGCCCTGGTGTATCCAGCTGCTCCGGATC 597
Db 541 ATGATGATGAATGCTGCGCTACAGCGCCCTGGTGTATCCAGCTGCTCCGGATC 597

RESULT 6

US-08-927-597-21
; Sequence 21, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..720
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..717
US-08-927-597-21

Query Match 93.9%; Score 597; DB 3; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.7e-147;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGGTCAATCATACCCCTTACATGCGGCTTCGCCACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAGGTCAATCATACCCCTTACATGCGGCTTCGCCACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGTT 120
Db 61 CCGCTCGTGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGTT 120
QY 121 CTGGAGGACGGGTGAACTATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
Db 121 CTGGAGGACGGGTGAACTATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGCTTCAGCTTCGGCTTATGAAGTGCACAACGTG 240
Db 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGCTTCAGCTTCGGCTTATGAAGTGCACAACGTG 240
QY 241 TCCGGATGTACCATGTACGAAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG 300
Db 241 TCCGGATGTACCATGTACGAAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATCATGACACACCCCGGGTGGCTGCGCTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGGCTGCGCTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCCAGGCTCCCAACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCCAGGCTCCCAACACGACA 420
QY 421 ATACGAGCCACGTGCAATTCAGCTGTTTCAACATCTCGCTCGCGGATGAGACGGTG 480
Db 421 ATACGAGCCACGTGCAATTCAGCTGTTTCAACATCTCGCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTCAGCTGTTTCAACATCTCGCTCGCGGATGAGACGGTG 540
Db 481 CAGGACTGCAATTCAGCTGTTTCAACATCTCGCTCGCGGATGAGACGGTG 540
QY 541 ATGATGATGAATGCTGCGCTACAAACGGCCCTGGTGTATCCAGCTGCTCCGGATC 597
Db 541 ATGATGATGAATGCTGCGCTACAAACGGCCCTGGTGTATCCAGCTGCTCCGGATC 597

RESULT 7

US-08-612-973-23
; Sequence 23, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100

us-09-899-303a-27.rn1

Mon Dec 22 13:28:44 2003

```

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..555
; US-08-612-973-23

Query Match      87.4%; Score 556; DB 3; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.4e-136;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGTTGGGTAAGGTGATCATGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGGTACATT 60
Db      1 ATGTTGGGTAAGGTGATCATGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGGTACATT 60
QY      61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGCTCCGGGTT 120
Db      61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGCTCCGGGTT 120
QY      121 CTGGAGGACGGGTGAACATGATGCAACAGGGAATTTGCCGGGTGCTTCTCTATCTTC 180
Db      121 CTGGAGGACGGGTGAACATGATGCAACAGGGAATTTGCCGGGTGCTTCTCTATCTTC 180
QY      181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG 240
Db      181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG 240
QY      241 TCCGGGATGTACCATGTGTCAGTCCGAACTCAAGCACTCAAGCACTCAAGCACTCAAGCACT 300
Db      241 TCCGGGATGTACCATGTGTCAGTCCGAACTCAAGCACTCAAGCACTCAAGCACTCAAGCACT 300
QY      301 GACATGATCATGCACACCCCCCGGGTCCGTCCTCGCTTCCAGCTTCGGGAGAACAACTCTTCCCGC 360
Db      301 GACATGATCATGCACACCCCCCGGGTCCGTCCTCGCTTCCAGCTTCGGGAGAACAACTCTTCCCGC 360
QY      361 TGCTGGGTAGGCTCACCCCCCAGCTCGAGCTAGGAAACGCCAGCGTCCCGGAT 420
Db      361 TGCTGGGTAGGCTCACCCCCCAGCTCGAGCTAGGAAACGCCAGCGTCCCGGAT 420
QY      421 ATACGACGCCACGTCGATTCACGATTCACCATCTCGCTCGCGCATGAGCGGTG 480
Db      421 ATACGACGCCACGTCGATTCACGATTCACCATCTCGCTCGCGCATGAGCGGTG 480
QY      481 CAGGACTGCAATTCGCTCAATCTATCCCGGCCACATAAGCGGTTCACCGTATGGCTTGGGAT 540
Db      481 CAGGACTGCAATTCGCTCAATCTATCCCGGCCACATAAGCGGTTCACCGTATGGCTTGGGAT 540
QY      541 ATGATGATGAACCTGGT 556
Db      541 ATGATGATGAACCTGGT 556

RESULT 8
US-08-927-597-23
; Sequence 23, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

```

```

; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 23:
; LENGTH: 561 base pairs
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..555
; US-08-927-597-23

Query Match      87.4%; Score 556; DB 3; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.4e-136;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGTTGGGTAAGGTGATCATGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGGTACATT 60
Db      1 ATGTTGGGTAAGGTGATCATGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGGTACATT 60
QY      61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGCTCCGGGTT 120
Db      61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGCTCCGGGTT 120
QY      121 CTGGAGGACGGGTGAACATGATGCAACAGGGAATTTGCCGGGTGCTTCTCTATCTTC 180
Db      121 CTGGAGGACGGGTGAACATGATGCAACAGGGAATTTGCCGGGTGCTTCTCTATCTTC 180
QY      181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG 240
Db      181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTTG 240
QY      241 TCCGGGATGTACCATGTGTCAGTCCGAACTCAAGCACTCAAGCACTCAAGCACTCAAGCACT 300
Db      241 TCCGGGATGTACCATGTGTCAGTCCGAACTCAAGCACTCAAGCACTCAAGCACTCAAGCACT 300
QY      301 GACATGATCATGCACACCCCCCGGGTCCGTCCTCGCTTCCAGCTTCGGGAGAACAACTCTTCCCGC 360
Db      301 GACATGATCATGCACACCCCCCGGGTCCGTCCTCGCTTCCAGCTTCGGGAGAACAACTCTTCCCGC 360
QY      361 TGCTGGGTAGGCTCACCCCCCAGCTCGAGCTAGGAAACGCCAGCGTCCCGGAT 420

```


Db 361 TGCTGGGTAGCGCTCACCCCAAGCTCGCAGCTAGGAACGCCAGCGTCCCCACACGACA 420
Qy 421 ATAGACGCCACGTCGATTCAGCTTCCAGCTGTTACCATCTCGCCCTCGCGCGCATGAGACGGTG 480
Db 421 ATAGACGCCACGTCGATTCAGCTTCCAGCTGTTACCATCTCGCCCTCGCGCGCATGAGACGGTG 480
Qy 481 CAGGACTCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Qy 541 ATGATGATGAAGTGGT 556
Db 541 ATGATGATGAAGTGGT 556

RESULT 9
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789

US-08-612-973-5
Query Match 81.0%; Score 515; DB 3; Length 795;
Best Local Similarity 89.2%; Pred. No. 8.1e-126;
Matches 597; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
Qy 1 ATGTTGGGTAAAGTATCATGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAAGTATCATGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
Qy 61 CCGCTCGTGGCGCGCCCTAGGGGGCGCTCCAGAGGCGCTTGGGGCATGGGGTTCGGGGTT 120
Db 61 CCGCTCGTGGCGCGCCCTAGGGGGCGCTCCAGAGGCGCTTGGGGCATGGGGTTCGGGGTT 120
Qy 121 CTGAGAGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTGTCTTTCTCTATCTTC 180
Db 121 CTGAGAGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTGTCTTTCTCTATCTTC 180
Qy 181 CTCTGGCTTTGCTGCTCTGACCGTTCCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 240
Db 181 CTCTGGCTTTGCTGCTCTGACCGTTCCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 240
Qy 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATTTGTATGAGCAGCG 300
Db 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATTTGTATGAGCAGCG 300
Qy 301 GACATGATCATGCACACCCCGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 GACATGATCATGCACACCCCGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGCTGGGTAGCGCTCACCCCGGCTCGCAGCTAGGACGCGTCCGCCACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCCGGCTCGCAGCTAGGACGCGTCCGCCACACGACA 420
Qy 421 ATACGACGCCACGTCGAT-----TCCAGCTGTTTCCGCTATGATGATGATGATGATG 480
Db 421 ATACGACGCCACGTCGAT-----TCCAGCTGTTTCCGCTATGATGATGATGATGATG 480
Qy 439 -----TCCAGCTGTTTCCAGCTGTTTCCGCTATGATGATGATGATGATGATGATG 468
Db 481 GGGGACCTCTCGGATCTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Qy 469 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTACCGGT 528
Db 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTACCGGT 600
Qy 529 ATGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 588
Db 601 ATGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 589 CTCGGGATC 597
Db 661 CTCGGGATC 669

RESULT 10
US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:

529 ATGGCTTGGATATCATGATGATGCTGCTGCTACACGGCCCTGGTGGTATCGAGCTG 588
501 ATGGCTTGGATATCATGATGATGCTGCTGCTACACGGCCCTGGTGGTATCGAGCTG 660
589 CTCGGGATC 597
661 CTCGGGATC 669

RESULT 11
US-08-612-973-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-612-973-47

Query Match 80.5%; Score 512; DB 3; Length 2082;
Best Local Similarity 89.2%; Pred. No. 6.4e-125;
Matches 594; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

4 TTGGGTAAGTATCATGATGATGCTGCTGCTACACGGCCCTGGTGGTATCGAGCTG 63
4 TTGGGTAAGTATCATGATGATGCTGCTGCTACACGGCCCTGGTGGTATCGAGCTG 63
64 CTGCTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGGTTCG 123
64 CTGCTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGGTTCG 123

APPLICATION NUMBER: US/08/927,597
FILING DATE: 11-MAR-1996
PRIOR APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789
US-08-927-597-5

Query Match 81.0%; Score 515; DB 3; Length 795;
Best Local Similarity 89.2%; Pred. No. 8.1e-126;
Matches 597; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

1 ATGTTGGTAAGTATCATGATGATGCTGCTGCTACACGGCCCTGGTGGTATCGAGCTG 60
1 ATGTTGGTAAGTATCATGATGATGCTGCTGCTACACGGCCCTGGTGGTATCGAGCTG 60
61 CCGCTCTCGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGGTTCGGGT 120
61 CCGCTCTCGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGGTTCGGGT 120
121 CTGAGAGCGGTGAATATGCAACAGGGAATTTGCCGGTTGCTTTCTCTATCTTTC 180
121 CTGAGAGCGGTGAATATGCAACAGGGAATTTGCCGGTTGCTTTCTCTATCTTTC 180
181 CTCTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
181 CTCTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
241 TCCGGGATGTACCATGTACGAAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
241 TCCGGGATGTACCATGTACGAAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
301 GACATGATCATGCAACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
301 GACATGATCATGCAACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
361 TGCTGGGTAGCGCTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
361 TGCTGGGTAGCGCTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
421 ATACGAGCGGATGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
421 ATACGAGCGGATGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
439 -----TCCCGAGCTGTTTACCATCTCGCTCGCGG 468
481 GGGGACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
469 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAACGGGTG 528
541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAACGGGTG 600

Db 664 CGGATC 669

|||||

RESULT 13

US-08-612-973-49

; Sequence 49, Application US/08612973

; Patent No. 6150134

; GENERAL INFORMATION:

; APPLICANT: MAERTENS, GEERT

; APPLICANT: BOSMAN, FONS

; APPLICANT: DE MARTYNOFF, GUY

; APPLICANT: BUYSE, MARIE-ANGE

; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHYE P.C.

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION NUMBER: US/08/612,973

; FILING DATE: 11-MAR-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: BYRNE, THOMAS E.

; REGISTRATION NUMBER: 32,205

; REFERENCE/DOCKET NUMBER: 1487-10

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 49:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2433 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..2430

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 1..2427

US-08-612-973-49

Query Match 80.5%; Score 512; DB 3; Length 2433;

Best Local Similarity 89.2%; Pred. No. 6.6e-125;

Matches 594; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 4 TTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATTCCG 63

Db 355 TTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATTCCG 414

QY 64 CTCGTGGCGCCCGCTAGCGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGGTTCTG 123

Db 415 CTCGTGGCGCCCGCTAGCGGCGCTGCGAGGCGCTGGCGCATGGCGTCCGGGTTCTG 474

QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTCTC 183

Db 475 GAGGACGGGTGAATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTCTC 534

QY 184 TTGGCTTTGCTGTCGTCTGACCGGTTCCAGCTTCGGCTTATGAAGTGGCAACGTGTC 243

Db 535 TTGGCTTTGCTGTCGTCTGACCGGTTCCAGCTTCGGCTTATGAAGTGGCAACGTGTC 594

QY 244 GGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAGCGAC 303

Db 595 GGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAGCGAC 654

QY 304 ATGATCATGCACACCGCGGCTGCGCTGCGTTCGGGAGAACAACTCTTCCGCTGC 363

Db 655 ATGATCATGCACACCGCGGCTGCGCTGCGTTCGGGAGAACAACTCTTCCGCTGC 714

QY 364 TGGGTAGCGCTCACCGCCACGCTCGCAGCTAGAAAGCGCAGGTCCTCCACACGCAATA 423

Db 715 TGGGTAGCGCTCACCGCCACGCTCGCAGCTAGAAAGCGCAGGTCCTCCACACGCAATA 774

QY 424 CGAGCCACGTCGAT-----TCCAGCTGTTTCCACCATTCGCTCGCGGCTATG 438

Db 775 CGAGCCACGTCGATTTGCTGTTGGGCGGCTGCTTTCTGTCGCTATGACGTGGG 834

QY 439 -----TCCAGCTGTTTCCACCATTCGCTCGCGGCTATG 471

Db 835 GACCTCGGGATCTGCTTCTCTCCAGCTTCCACCATTCGCTCGCGGCTATG 894

QY 472 GAGACGCTGAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTACCGTATG 531

Db 895 GAGACGCTGAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTACCGTATG 954

QY 532 GCTTGGGATATGATGAACTGCTGCTTACAAACGGGCTTGGTGTATCGCAGCTGCTC 591

Db 955 GCTTGGGATATGATGAACTGCTGCTTACAAACGGGCTTGGTGTATCGCAGCTGCTC 1014

QY 592 CGGATC 597

Db 1015 CGGATC 1020

RESULT 14

US-08-927-597-49

; Sequence 49, Application US/08927597

; Patent No. 6245503

; GENERAL INFORMATION:

; APPLICANT: MAERTENS, GEERT

; APPLICANT: BOSMAN, FONS

; APPLICANT: DE MARTYNOFF, GUY

; APPLICANT: BUYSE, MARIE-ANGE

; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/927,597

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/612,973

; FILING DATE: 11-MAR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: BYRNE, THOMAS E.

; REGISTRATION NUMBER: 32,205

; REFERENCE/DOCKET NUMBER: 1487-10

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000


```

; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 2433 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;
; MOLECULE TYPE: cDNA
;
; HYPOTHETICAL: NO
;
; ANTI-SENSE: NO
;
; FEATURE:
;
;     NAME/KEY: CDS
;     LOCATION: 1..2430
;
;     FEATURE:
;
;     NAME/KEY: mat_peptide
;     LOCATION: 1..2427
;
; US-08-927-597-49

```

Query Match	80.5%	Score 512	DB 3	Length 2433
Best Local Similarity	89.2%	Pred. No. 6.6e-125		
Matches 594	Conservative 0	Mismatches 0	Indels 72	Gaps 1
QY	4	TTGGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGG	63	
Db	355	TTGGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGG	414	
QY	64	CTCGTCGCGCGCCCCCTAGGGGGCGCTGCCACAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	123	
Db	415	CTCGTCGCGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	474	
QY	124	GAGGACGGCGTGAACATATGCAACAGGGAAATTTGCCCCGGTTGCTCTTCTATCTCTATCTCCTC	183	
Db	475	GAGGACGGCGTGAACATATGCAACAGGGAAATTTGCCCCGGTTGCTCTTCTATCTCTATCTCCTC	534	
QY	184	TTGGCTTTTGCTGTCCTCTGACCGTTCCAGCTTTCGGCTTATGAAGTCCGCAACGTTGTC	243	
Db	535	TTGGCTTTTGCTGTCCTCTGACCGTTCCAGCTTTCGGCTTATGAAGTCCGCAACGTTGTC	594	
QY	244	GGGATGTACCATTGTACAGCAACGACTGTCTCAAACCTCAAGCATTTGTGTATGAGGCAGCGGAC	303	
Db	595	GGGATGTACCATTGTACAGCAACGACTGTCTCAAACCTCAAGCATTTGTGTATGAGGCAGCGGAC	654	
QY	304	ATGATCATGCAACACCCCGGGTGGTGCGCTTCGGGAGAACAACTCTTTCGCGTGC	363	
Db	655	ATGATCATGCAACACCCCGGGTGGTGCGCTTCGGGAGAACAACTCTTTCGCGTGC	714	
QY	364	TGGGTAGCGGTACACCCACGCTCGCAGCTAGGAACGCCAGCGTCCCAACACACACAATA	423	
Db	715	TGGGTAGCGGTACACCCACGCTCGCAGCTAGGAACGCCAGCGTCCCAACACACACAATA	774	
QY	424	CGACGCCACGTCGAT	438	
Db	775	CGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCGGCTATGTACGTGGG	834	
QY	439	-----TCCAGCTGTTTCAACATCTCGCTTCGCGCGGAT	471	
Db	835	GACCTCTGCGGATCTGTCTTCTCCGCTCTCCAGCTGTTTCAACATCTCGCTTCGCGCGGAT	894	
QY	472	GAGACGGTCGAGGACTCGAATTCGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATG	531	
Db	895	GAGACGGTCGAGGACTCGAATTCGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATG	954	
QY	532	GCTTTGGGATATGATGATGAATCGTTCGCTCAACACGGCCCTTGGTGGTATCGACGTCGTC	591	
Db	955	GCTTTGGGATATGATGATGAATCGTTCGCTCAACACGGCCCTTGGTGGTATCGACGTCGTC	1014	
QY	592	CGGATC	597	
Db	1015	CGGATC	1020	

RESULT 15
US-08-612-973-7

```

: Sequence 7, Application US/08612973
: Patent No. 6150134
: GENERAL INFORMATION:
: APPLICANT: MAERTENS, GEERT
: APPLICANT: BOSMAN, FONS
: APPLICANT: DE MARTYNOFF, GUY
: APPLICANT: BUYSSE, MARIE-ANGE
: TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
: TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
: NUMBER OF SEQUENCES: 111
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHYE P.C.
: STREET: 1100 NORTH GLEBE ROAD
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: U.S.A.
: ZIP: 22201-4714
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/612,973
: FILING DATE: 11-MAR-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: BYRNE, THOMAS E.
: REGISTRATION NUMBER: 32,205
: REFERENCE/DOCKET NUMBER: 1487-10
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 816-4000
: TELEFAX: (703) 816-4100
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 633 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..630
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 1..627
: US-08-612-973-7

```

Query Match	71.8%; Score 456.4; DB 3; Length 633;
Best Local Similarity	86.8%; Pred. No. 1.7e-110;
Matches 545; Conservative	0; Mismatches 11; Indels 72; Gaps 1;
Qy	1 ATGTTGGTAAAGTCATCGATACCCATTACATGCGGCTTCGCCACCTCGTGGGGTACATT 60
Db	1 ATGTTGGTAAAGTCATCGATACCCATTACGTTGCGGGCTTCGCCACCTCATGGGGTACATT 60
Qy	61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCAGAGGCCCTGCGCGATGCGCTCCGGGTT 120
Db	61 CCGCTCGTCGGCGCCCCCTTAGGGGGTGTGCGCAGAGCCCTGCGCGATGCGCTCCGGGTT 120
Qy	121 CTGAGGACGGCGTGAACACTATGCAACAGGGAAATTTGGCCGGTTGCTCTTCTCTATCTTC 180
Db	121 CTGGAAGACGGCGTGAACACTATATGCAACAGGGAAATTTGGCTGGTTGCTCTTCTCTATCTTC 180
Qy	181 CTCCTTGGCTTTGCTGTCTCTGTGACCGTTTCCAGCTTTCGCGTTATGAAGTGCGCACAGTG 240
Db	181 CTCCTTGGCTTTACTGTCTCTGTGACCAATTCACAGTTCGCGTTATGAGGTGCGCAACAGTG 240
Qy	241 TCCGGGATGACCAATGTCACGAACGACTGTCCAACTCCAAAGCTATGATAGGCGACGC 300
Db	241 TCCGGGATGACCAATGTCACGAACGACTGTCCAACTCAAGCAATTTGTTATGAGGCGACGC 300

Mon Dec 22 13:28:44 2003

QY	301	GACATGATCATGCACACCCCGGGTGGTGCCTGGTTCGGGAGAACTCTTCCCGC	360
Db			
QY	301	GACATGATCATGCACACACCCCGGGTGGTGCCTGGTTCGGGAGAACTCTTCCCGC	360
Db			
QY	361	TGCTGGGTAGCGCTCACCCCGCACCGCTCGCAGCTAGGAACGCCAGCGTCCCGCACCGACA	420
Db			
QY	361	TGCTGGGTAGCGCTCACCCCGCACCGCTCGCAGCTAGGAACGCCAGCGTCCCGCACCGACA	420
Db			
QY	421	ATACGACGCCACGTCGAT-----	438
Db			
QY	421	ATACGACGCCACGTCGATTTGCTGTGGGGCGGCTGCTTTCTGTTCGCTATGTACGTG	480
Db			
QY	439	-----TCCGAGCTGTTACCATCTCGCCTCGCGG	468
Db			
QY	481	GGGGATCTCTGGGATCTGTCTTCTCGTCTCCGAGCTTTCACCATCTCGCCTCGCGG	540
Db			
QY	469	CATGACCGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT	528
Db			
QY	541	CATGACCGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT	600
Db			
QY	529	ATGGCTTGGGATATGATGATGACTGGT	556
Db			
QY	601	ATGGCTTGGGATATGATGATGACTGGT	628
Db			

Search completed: December 20, 2003, 07:03:15
Job time : 48.8681 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 19, 2003, 16:55:48 ; Search time 2385.36 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-29
Perfect score: 630
Sequence: 1 ATGGGTAAAGTCATCGATAC.....TGATGATGAACCTGGTATATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgtgo_hum.*
- 40: em_hgtgo_mus.*
- 41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	630	100.0	630	6	A48691	A48691 Sequence 29
2	630	100.0	630	6	AR157340	AR157340 Sequence
3	630	100.0	630	6	AX452778	AX452778 Sequence
4	630	100.0	630	6	AX685030	AX685030 Sequence
5	587.2	93.2	1546	14	HPCSTRUCTC	L12355 Hepatitis C
6	585.6	93.0	1786	14	HPCNZLICR	D14305 Hepatitis C
7	585.6	93.0	9390	14	HVCVENS1	X76918 Hepatitis C
8	585.6	93.0	9456	14	HPCCEGS	D17763 Hepatitis C
9	582.4	92.4	1786	14	HPCUS114CE	D14309 Hepatitis C
10	576	91.4	1786	14	HPCTH85CE	D14307 Hepatitis C
11	576	91.4	9425	14	AF046866	AF046866 Hepatitis C
12	576	91.4	9454	14	HPCK3A	D28917 Hepatitis C
13	574.4	91.2	1786	14	HPCHEM26CE	D14311 Hepatitis C
14	539.4	85.6	541	6	A40621	A40621 Sequence 21
15	539.4	85.6	541	6	AX031599	AX031599 Sequence
16	539.4	85.6	541	6	AX031869	AX031869 Sequence
17	539.4	85.6	541	6	AX032139	AX032139 Sequence
18	539.4	85.6	541	6	BD172134	BD172134 New seque
19	539.4	85.6	541	14	HPCCOREED	D14599 Hepatitis C
20	533	84.6	541	6	A40619	A40619 Sequence 19
21	533	84.6	541	6	AX031597	AX031597 Sequence
22	533	84.6	541	6	AX031867	AX031867 Sequence
23	533	84.6	541	6	AX032137	AX032137 Sequence
24	533	84.6	541	6	BD172133	BD172133 New seque
25	504.2	80.0	541	6	A40613	A40613 Sequence 13
26	504.2	80.0	541	6	A40617	A40617 Sequence 17
27	504.2	80.0	541	6	AX031591	AX031591 Sequence
28	504.2	80.0	541	6	AX031595	AX031595 Sequence
29	504.2	80.0	541	6	AX031861	AX031861 Sequence
30	504.2	80.0	541	6	AX031865	AX031865 Sequence
31	504.2	80.0	541	6	AX032131	AX032131 Sequence
32	504.2	80.0	541	6	AX032135	AX032135 Sequence
33	504.2	80.0	541	6	BD172130	BD172130 New seque
34	504.2	80.0	541	6	BD172132	BD172132 New seque
35	504.2	80.0	541	14	HPCCOREEH	D14603 Hepatitis C
36	502.6	79.8	541	6	A40615	A40615 Sequence 15
37	502.6	79.8	541	6	AX031593	AX031593 Sequence
38	502.6	79.8	541	6	AX031863	AX031863 Sequence
39	502.6	79.8	541	6	AX032133	AX032133 Sequence
40	502.6	79.8	541	6	BD172131	BD172131 New seque
41	499.4	79.3	541	6	A40623	A40623 Sequence 23
42	499.4	79.3	541	6	A40625	A40625 Sequence 25
43	499.4	79.3	541	6	A40627	A40627 Sequence 27
44	499.4	79.3	541	6	AX031601	AX031601 Sequence
45	499.4	79.3	541	6	AX031603	AX031603 Sequence

ALIGNMENTS

RESULT 1

A48691
LOCUS A48691 Sequence 29 from Patent WO9604385. 630 bp DNA linear PAT 07-MAR-1997
DEFINITION A48691
ACCESSION A48691
VERSION A48691.1 GI:2302404
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 630)
AUTHORS Maertens, G., Bosman, F., De M.G. and Buyse, M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 29 15-FEB-1996;

Mon Dec 22 13:28:45 2003

INNOGENETICS NV (BE)		AR157340	Sequence 29 from patent US 6245503.	630 bp	DNA	linear	PAT 17-OCT-2001
COMMENT		Other publication CA 2172273 960215	AR157340				
Other publication AU 3382495 960304.		AR157340.1	GI:16218274				
FEATURES		Location/Qualifiers	1..630				
source		/organism="unidentified"					
		/mol_type="genomic DNA"					
		/db_xref="taxon:32644"					
CDS		1..627					
		/note="unnamed protein product"					
		/codon_start=1					
		/protein_id="CAA03138.1"					
		/db_xref="GI:2302405"					
		/translation="MGKVIDTLTCGFADLMGVPLVAGPVGVARALAHGVRALEDGI NFAIENLPCGSFSIFLLALPCLLHPASLEWRNTSLGYLVTNDCSNSIVYEADDVI LHPGICPCVQDGNSTCTWPTVAVKYVGATASIRSHVDLLVGAATMCSALYVG DMCGAVFLVGOAFTFRPRRHQIVQTCNSLPHLSGHRMAWDMMNW"					
mat_peptide		1..624					
BASE COUNT		127 a 175 c 168 g 160 t					
ORIGIN							
Query Match		100.0%; Score 630; DB 6; Length 630;					
Best Local Similarity		100.0%; Pred. No. 4.5e-165;					
Matches		630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	ATGGGTAAGTTCATCGATACCCCTTACGTGGGATTCGCGATCTCATGGGGTACATCCCG	60				
Db	1	ATGGGTAAGTTCATCGATACCCCTTACGTGGGATTCGCGATCTCATGGGGTACATCCCG	60				
Qy	61	CTCGTCGGCGCTCCCGTAGGAGCGGTGCGAAGAGCCCTTGGCGATCGCGTGGGGCCCTT	120				
Db	61	CTCGTCGGCGCTCCCGTAGGAGCGGTGCGAAGAGCCCTTGGCGATCGCGTGGGGCCCTT	120				
Qy	121	GAAGACGGGATAAATTTGCAACAGAGGAAATTTGCCGGTGTCTTCTATTTTCCTT	180				
Db	121	GAAGACGGGATAAATTTGCAACAGAGGAAATTTGCCGGTGTCTTCTATTTTCCTT	180				
Qy	181	CTCGCTCTGTTCTCTGCTTAATTCACAGAGTGTCTAGAGTGGCGGAATACGTCT	240				
Db	181	CTCGCTCTGTTCTCTGCTTAATTCACAGAGTGTCTAGAGTGGCGGAATACGTCT	240				
Qy	241	GGCCTCTATGCTTCAACAGCTGTTCCTAATAGCAGTGTGTAGAGTGGCGGATGAC	300				
Db	241	GGCCTCTATGCTTCAACAGCTGTTCCTAATAGCAGTGTGTAGAGTGGCGGATGAC	300				
Qy	301	GTTATCTGACACACCCGGCTGCATACCTTGTCTCCAGGACGGCAATACATCCAGTGC	360				
Db	301	GTTATCTGACACACCCGGCTGCATACCTTGTCTCCAGGACGGCAATACATCCAGTGC	360				
Qy	361	TGGACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTGGAGCAACCCGCTTCGATA	420				
Db	361	TGGACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTGGAGCAACCCGCTTCGATA	420				
Qy	421	CGCAGTCATGTGGACCTTATAGTGGCGCGCCAGATGTCTCTGCGCTCTACGTGGGT	480				
Db	421	CGCAGTCATGTGGACCTTATAGTGGCGCGCCAGATGTCTCTGCGCTCTACGTGGGT	480				
Qy	481	GACATGTGGGGCTGTCTTCTCGTGGGACAAAGCCCTTCAAGTTCAGACCTTCGTGGCAT	540				
Db	481	GACATGTGGGGCTGTCTTCTCGTGGGACAAAGCCCTTCAAGTTCAGACCTTCGTGGCAT	540				
Qy	541	CAAAACGGTCCAGACCTGTAACTCTCGCTGTACCCAGGCCATCTTTTCAGGACATCGAATG	600				
Db	541	CAAAACGGTCCAGACCTGTAACTCTCGCTGTACCCAGGCCATCTTTTCAGGACATCGAATG	600				
Qy	601	GCTTTGGGATATGATGATGAACCTGGTAATAG	630				
Db	601	GCTTTGGGATATGATGATGAACCTGGTAATAG	630				
RESULT 3		AX452778	630 bp	DNA	linear	PAT 06-JUL-2001	
LOCUS		AX452778					
DEFINITION		Sequence 29 from Patent EP1211315.					
ACCESSION		AX452778					


```

VERSION      AX452778.1  GI:21712463
KEYWORDS     Hepatitis C virus
SOURCE       Hepatitis C virus
ORGANISM     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE        Recombinant vectors for producing hcv envelope proteins
JOURNAL      Patent: EP 1211315-A 29 05-JUN-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
              1..630
              /organism="Hepatitis C virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:11103"
              1..627
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAD38076.1"
              /db_xref="GI:21712464"
              /translation="MGKVIDTLTCGFADLMGYIPLVGAPVGVARALAHGVRALDGI
              NFATNLPGCSFISFLALFSLIHPAASLEWRNTSGLYVLNDSCNSIIVYAADVI
              LHTPGCIPCVODGNTSTCWTPVTPVAVKYVGATTASIRSHVDLLVGATMCSALYVG
              DMCAGVFLVQAFTPRPRHQVTQCNCSLYPHLSGHRMADMMNNW"
              mat_peptide 1..624
              /product="unnamed"
BASE COUNT   127 a 175 c 168 g 160 t
ORIGIN
Query Match 100.0%; Score 630; DB 6; Length 630;
Best Local Similarity 100.0%; Pred. No. 4.5e-165;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGTAAAGTCATCGATACCCCTTACGTCGGGATTCGCGATCTCATGGGGTACATCCCG 60
DB 1 ATGGGTAAAGTCATCGATACCCCTTACGTCGGGATTCGCGATCTCATGGGGTACATCCCG 60
QY 61 CTCGTGGCGCTCCCGTAGGAGCGCTGCGAAGAGCCCTTGCGCATGGCGGCGGATACGCT 120
DB 61 CTCGTGGCGCTCCCGTAGGAGCGCTGCGAAGAGCCCTTGCGCATGGCGGCGGATACGCT 120
QY 121 GAAGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTGCTCTTTTCTATTTCCTT 180
DB 121 GAAGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTGCTCTTTTCTATTTCCTT 180
QY 181 CTCGCTCTGTCTCTTCTTAAATTCATCCAGCAGTAGTCTAGAGTGGCGGAATACGCT 240
DB 181 CTCGCTCTGTCTCTTCTTAAATTCATCCAGCAGTAGTCTAGAGTGGCGGAATACGCT 240
QY 241 GGCCTCTATGTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAC 300
DB 241 GGCCTCTATGTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAC 300
QY 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCAGTGC 360
DB 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCAGTGC 360
QY 361 TGGACCCCAAGTGAACCTTACAGTGGCAGTCAAGTACGTGGAGCAACACCCGCTTCGATA 420
DB 361 TGGACCCCAAGTGAACCTTACAGTGGCAGTCAAGTACGTGGAGCAACACCCGCTTCGATA 420
QY 421 CGCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT 480
DB 421 CGCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT 480
QY 481 GACATGTGTGGGGTGTCTTCTCTGTGGGACAAGCCTTACAGTTCAGACCTCGTGGCCAT 540
DB 481 GACATGTGTGGGGTGTCTTCTCTGTGGGACAAGCCTTACAGTTCAGACCTCGTGGCCAT 540
QY 541 CAAACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCCATCTTTTCAGGACATCGAATG 600
DB 541 CAAACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCCATCTTTTCAGGACATCGAATG 600

RESULT 4
AX685030
LOCUS       AX685030
DEFINITION Sequence 29 from Patent WO02055548.
ACCESSION  AX685030
VERSION     AX685030.1  GI:29371435
KEYWORDS    Hepatitis C virus
SOURCE      Hepatitis C virus
ORGANISM    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens,G., Bosman,F. and Buyse,M.A.
TITLE        Purified Hepatitis C Virus envelope proteins for diagnostic and
              therapeutic use
JOURNAL      Patent: WO 02055548-A 29 18-JUL-2002;
              INNOGENETICS N.V. (BE)
FEATURES     Location/Qualifiers
              1..630
              /organism="Hepatitis C virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:11103"
              1..627
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAD86529.1"
              /db_xref="GI:29371436"
              /translation="MGKVIDTLTCGFADLMGYIPLVGAPVGVARALAHGVRALDGI
              NFATNLPGCSFISFLALFSLIHPAASLEWRNTSGLYVLNDSCNSIIVYAADVI
              LHTPGCIPCVODGNTSTCWTPVTPVAVKYVGATTASIRSHVDLLVGATMCSALYVG
              DMCAGVFLVQAFTPRPRHQVTQCNCSLYPHLSGHRMADMMNNW"
              mat_peptide 1..624
              /product="unnamed"
BASE COUNT   127 a 175 c 168 g 160 t
ORIGIN
Query Match 100.0%; Score 630; DB 6; Length 630;
Best Local Similarity 100.0%; Pred. No. 4.5e-165;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGTAAAGTCATCGATACCCCTTACGTCGGGATTCGCGATCTCATGGGGTACATCCCG 60
DB 1 ATGGGTAAAGTCATCGATACCCCTTACGTCGGGATTCGCGATCTCATGGGGTACATCCCG 60
QY 61 CTCGTGGCGCTCCCGTAGGAGCGCTGCGAAGAGCCCTTGCGCATGGCGGCGGATACGCT 120
DB 61 CTCGTGGCGCTCCCGTAGGAGCGCTGCGAAGAGCCCTTGCGCATGGCGGCGGATACGCT 120
QY 121 GAAGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTGCTCTTTTCTATTTCCTT 180
DB 121 GAAGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTGCTCTTTTCTATTTCCTT 180
QY 181 CTCGCTCTGTCTCTTCTTAAATTCATCCAGCAGTAGTCTAGAGTGGCGGAATACGCT 240
DB 181 CTCGCTCTGTCTCTTCTTAAATTCATCCAGCAGTAGTCTAGAGTGGCGGAATACGCT 240
QY 241 GGCCTCTATGTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAC 300
DB 241 GGCCTCTATGTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAC 300
QY 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCAGTGC 360
DB 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCAGTGC 360
QY 361 TGGACCCCAAGTGAACCTTACAGTGGCAGTCAAGTACGTGGAGCAACACCCGCTTCGATA 420
DB 361 TGGACCCCAAGTGAACCTTACAGTGGCAGTCAAGTACGTGGAGCAACACCCGCTTCGATA 420
QY 421 CGCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT 480
DB 421 CGCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT 480
QY 481 GACATGTGTGGGGTGTCTTCTCTGTGGGACAAGCCTTACAGTTCAGACCTCGTGGCCAT 540
DB 481 GACATGTGTGGGGTGTCTTCTCTGTGGGACAAGCCTTACAGTTCAGACCTCGTGGCCAT 540
QY 541 CAAACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCCATCTTTTCAGGACATCGAATG 600
DB 541 CAAACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCCATCTTTTCAGGACATCGAATG 600
```


QY	421	CGCAGTCATGTGGACCTATTAGTGGCGCGCCACGATGTGCTCTGGCTCTACGTGGT	480
Db	421	CGCAGTCATGTGGACCTATTAGTGGCGCGCCACGATGTGCTCTGGCTCTACGTGGT	480
QY	481	GACATGTGTGGCGTCTCTCTGTGGGACAAAGCTTTCAGTTTCAGACCTCGTGGCAT	540
Db	481	GACATGTGTGGCGTCTCTCTGTGGGACAAAGCTTTCAGTTTCAGACCTCGTGGCAT	540
QY	541	CAAAAGGTCACACCTGTAACCTGCTGTGTACCCAGGCGCATCTTTTACGACATCGAATG	600
Db	541	CAAAAGGTCACACCTGTAACCTGCTGTGTACCCAGGCGCATCTTTTACGACATCGAATG	600
QY	601	GCTTGGATATGATGATGAACTGGTAATAG	630
Db	601	GCTTGGATATGATGATGAACTGGTAATAG	630
RESULT 5			
HPCTRUCTC		1546 bp ss-RNA linear	VRL 27-JUL-1994
LOCUS		Hepatitis C virus core and envelope proteins gene, 5' end of cds.	
DEFINITION			
ACCESSION		L12355	
VERSION		L12355.1	GI:410169
KEYWORDS		core protein; envelope protein; structural region.	
SOURCE		Hepatitis C virus	
ORGANISM		Hepatitis C virus	
REFERENCE		Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.	
AUTHORS		1 (sites)	
TITLE		Han, J.H., Shyamala, V., Richman, K.H., Brauer, M.J., Irvine, B., Urdes, M.S., Tekamp-Olson, P., Kuo, G., Choo, Q.L. and Houghton, M. Characterization of the terminal regions of hepatitis C viral RNA: Identification of conserved sequences in the 5' untranslated region and poly(A) tails at the 3' end	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 88 (5), 1711-1715	(1991)
MEDLINE		91156678	
PUBMED		1705704	
REFERENCE		2 (sites)	
AUTHORS		Choo, Q.-L., Richman, K., Han, J.H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, A., Barr, P.J., Weiner, A., Bradley, D.W., Kuo, G. and Houghton, M. Genetic organization and diversity of the hepatitis C virus	
TITLE		Genetic organization and diversity of the hepatitis C virus	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 88 (6), 2451-2455	(1991)
MEDLINE		91172826	
PUBMED		1848704	
REFERENCE		3 (bases 1 to 1546)	
AUTHORS		Li, J.S., Vitvitski, L., Tong, S.P. and Trepo, C.	
TITLE		Identification of the third major genotype of hepatitis C virus in France	
JOURNAL		Biochem. Biophys. Res. Commun. 199 (3), 1474-1481	(1994)
MEDLINE		94197744	
PUBMED		8147893	
COMMENT		Original source text: Hepatitis C virus cDNA to genomic RNA.	
FEATURES		Location/Qualifiers	
source		1..1546	
		/organism="Hepatitis C virus"	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
CDS			
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	
		/note="specific host (French)"	
		314..>1546	
		/standard_name="structural region"	
		/note="encodes core and envelope proteins"	
		/codon_start=1	
		/protein_id="AAA20155.1"	
		/db_xref="GI:410170"	
		/translation="MSTLPKPKQRTKRTIRRPQDVPPGGQIVGVYVLPKRPRL CYRATKTSQPRGRQPIPKARRSDGRSWAOPGPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPSRNIGKVIDLTLCGFADLMGVIPLVGPVGGVARALAHGVRALED GINFATNIPGCSFSIFLLALPSCLIHPAASLEWNTSGLYILTNDCNSNSIVYEAD VILHTPGCPVQDQSTCTWPTPTVAVRYGVGTASIRSHVDLLVGAATMCSALY VGMCGAVFLVQGAFTFRRHQTVQTCNSLYPGHLSGRHMDMMWNPAPGMVY AHLRLPQTLFDIIAGAHWGLLAGLAYYSMQGNWAKVAILIMVNFSGVDAGTRIIGGSA	
		/db_xref="taxon:11103"	
		/tissue type="serum"	


```

REFERENCE 2 (bases 1 to 1786)
AUTHORS Okamoto, H.
JOURNAL Unpublished
COMMENT Submitted (28-JAN-1993) to DDBJ by:
Hiroaki Okamoto
Immunology Division
Gichi Medical School
Kawachi
Tochigi 329-04
Japan
Phone: 0285-44-2111x3334
Fax: 0285-44-1557
FEATURES
source
1..1786
Location/Qualifiers
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="NZL1"
/db_xref="taxon:11103"
<1..277
/citation=[1]
278..>1786
/codon_start=1
/product="polyprotein precursor"
/protein_id="BAA03251.1"
/db_xref="GI:471114"
/translation="MSTLPKPKYKNTIRRPQDKFPGGQIVGGVTVLPRRPRRL
GVARTKTSERSQPRRQPIPKARRSEGRSWAQFYPMPLVNGCGWAGLLSPRG
SRPSWGNDRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRALD
GINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSLGYLTVNDCSNSIVYBADDD
VILHTPGVCVQDQNTSTCTWPTPTVAVRVGATTASIRSHVDLLVGAATMCSALY
VDMCGAVFLVQCATFTFRRHQTQVTCNCSLYPGHLSHRMANWMMNWSPAVGNV
AHVLKLPQTLFIAGAHWGLIAGLYSMQGNWAKVAIIMVMFSGVDAAETITGGTA
SRHTQAFAGLFDPQKQLQVKTNGSWHINSTALNCNCSINTGFIAGLYFYHKFNST
GCPQLSSCKPITFFRQGWGLTDANITGPSDDRPYCHYAPRPCDIIVPASSV"
278..850
/product="C"
51..1426
/product="E"
1427..>1786
/product="E2/NS1"
BASE COUNT 366 a 521 c 497 g 402 t
ORIGIN
mat_peptide
mat_peptide
mat_peptide
Query Match 93.0%; Score 585.6; DB 14; Length 1786;
Best Local Similarity 96.2%; Pred. No. 1.2e-152;
Matches 600; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 2 TGGTAAAGTCATCGATACCTTACGTGCGGATTGCGCGATCTCATGGGGTACATCCCGC 61
DB 633 TGGGTAAGTCATCGATACCTTACGTGCGGATTGCGCGACCTCATGGGGTACATCCCGC 692
QY 62 TCGTCGGGCTCCGTAGGAGCGTCGCAAGAGCCCTTCGCGATGCGGTAGGCGCCCTTG 121
DB 693 TCGTCGGGCTCCTGTAGAGGCGTCGCAAGAGCCCTTCGCGATGCGGTAGGCGCCCTTG 752
QY 122 AAGACGGGATAAATTTGCAACAGGAAATTCGCCGGTTCCTCTTTCTATTTTCCTTC 181
DB 753 AAGACGGGATAAATTTGCAACAGGAAATTCGCCGGTTCCTCTTTCTATTTTCCTTC 812
QY 182 TCGCTCTGTTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGTG 241
DB 813 TTGCTCTGTTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGTG 872
QY 242 GCCTCTATCTCTTACCAAGACTGTTCCCAATAGCAGTATTGTATACGAGGCGCGATGAGC 301
DB 873 GCCTCTATCTCTTACCAAGACTGTTCCCAATAGCAGTATTGTATAGGCGCGATGAGT 932
QY 302 TTATTCGCACACACCGCGCTGCATACCTTGTGTCCAGACGCGCAATACATCCACGTGCT 361
DB 933 TCATTCGCACACACCGCGCTGTGTACCTTGTGTCCAGACGCGCAATACATCTACGTGCT 992
QY 362 GGACCCCACTGACACCTTACAGTGGCAGTCAAGTACGTCGAGCAACACCGCTTCGATAC 421

```

```

Db 993 GGACCCCACTGACACCTTACAGTGGGAGTCAGGTACGTGGAGCAAACTACTGCTTCGATAC 1052
QY 422 GCAGTCATGTGACACCTTATAGTGGCGGGCCACGATGTGCTCTCGCTCTACGTGGGTG 481
Db 1053 GCAGTCATGTGACACCTTATAGTGGCGGGCCACGATGTGCTCTCGCTCTACGTGGGTG 1112
QY 482 ACATGTGTGGGCTGTCTTCTCTCGTGGGACACGCTTACGTTTACAGCTCTGCGCATC 541
Db 1113 ATATGTGTGGGCTGTCTTCTCTCGTGGGACACGCTTACGTTTACAGCTCTGCGCATC 1172
QY 542 AAACGGTCCAGACCTGTAACTGCTCGCTGACCCAGGCCATCTTTTCAGGACATCGAATCG 601
Db 1173 AAACGGTCCAGACCTGTAACTGCTCGCTGACCCAGGCCATCTTTTCAGGACATCGAATCG 1232
QY 602 CTTGGGATATGATGAACTGGT 625
Db 1233 CTTGGGATATGATGAACTGGT 1256

RESULT 7
LOCUS HVCENS1 9390 bp RNA linear VRL 17-JAN-1995
DEFINITION Hepatitis C virus genes for core, envelope and NS1 proteins.
ACCESSION X76918
VERSION X76918.1 GI:633201
KEYWORDS core protein; envelope protein; NS1 protein.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES: asRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 9390)
AUTHORS Seelig,R., Weber,P., Seeling,H.P., Ledger,N., Bottner,C. and
Renz,M.
TITLE Hepatitis C virus type V genome isolated from a patient in Germany
JOURNAL Unpublished
AUTHORS Seelig,R.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-1993) R. Seelig, Inst. of Immunol. and Mol.
Genetics, Kriegsstrasse 99, D- 76133 Karlsruhe, FRG
REFERENCE 3 (bases 1 to 9390)
AUTHORS Seelig,R.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-1995) R. Seelig, Inst. of Immunol. and Mol.
Genetics, Kriegsstrasse 99, D- 76133 Karlsruhe, FRG
COMMENT On Jan 24, 1995 this sequence version replaced gi:506489.
FEATURES
source
1..9390
/organism="Hepatitis C virus"
/mol_type="unassigned RNA"
/db_xref="taxon:11103"
1..2267
/organism="Hepatitis C virus"
/mol_type="unassigned RNA"
/sub_species="type V-D (Germany)"
/db_xref="taxon:11103"
/maps="-297 to 1970 nucleotide number"
/tissue_type="serum"
299..9364
/note="viral coding region"
/codon_start=1
/protein_id="CAA54244.1"
/db_xref="GI:633202"
/db_xref="SPTREMBL:Q68870"
/translation="MSTLPKPKYKNTIRRPQDKFPGGQIVGGVTVLPRRPRRL
GVARTKTSERSQPRRQPIPKARRSEGRSWAQFYPMPLVNGCGWAGLLSPRG
SRPSWGNDRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRALD
GINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSLGYLTVNDCSNSIVYBADDD
VILHTPGVCVQDQNTSTCTWPTPTVAVRVGATTASIRSHVDLLVGAATMCSALY
VDMCGAVFLVQCATFTFRRHQTQVTCNCSLYPGHLSHRMANWMMNWSPAVGNV
SHVLKLPQTLFIAGAHWGLIAGLYSMQGNWAKVAIIMVMFSGVDAAETITGGTA
AHGVTLLTSLFSSGPKQLQVKTNGSWHINSTALNCNCSINTGFIAGLYFYHKFNST
GCPQLSSCKPITFFRQGWGLTDANITGPSDDRPYCHYAPRPCDIIVPASSV"
CDS

```


CFTSPVVVGGTTDRKVPYTNWGENESDVFLLSESRPPSGRWFGCAWNNSTGFLKTCG
APPCNIYGGGNNPNHSHLFCPTDCFRKPHQDATYSPSCGAGPWLTPRCMVDYFYRLMHY
PCTVNPFLFKVMEVGFHRRFTTAACNWTGRCERNICEDRRSEQHPLHSTHTEALILP
CSTFMPMLQTLHLHQNLIVDQVLYGVSGMGVWALKBEFVILFILLADARVCVA
LWLMITISQASBALENLIVTKAVAAAGTGHIGHYLVAFCAAWHVKHGLPLVTVSYLTC
LWSTALLVLLPQRAYAWSGEDSVLITLSSLPYSLIFDITKLILAVLGPYLIQASITA
CRCEALQVWPYPILLARQVGFQOMILSCGLPVSARLBSREGLVHGPADDAWAGLKR
TPYFVRAHVLRLCMLVIRSMGAGTKAACDILCGLPVSARLBSREGLVHGPADDAWAGLKR
LAVAPEVILFMEIKVITWAGMTAAACDILCGLPVSARLBSREGLVHGPADDAWAGLKR
LLAAPTAAQOTGRGLGTVITGLTRGNVVTVEQVLSSTATQTFGLTGVGVVMTVI
HGAQSRTLAGAKHPALOMYTNVDOLVWGPAPGACSLPCCAGDZKIPVAAVYGVNVLV
PARRGDSRSLSPRLPCLAKSGSGGVMPCSHVAGIFRAVCTRGVAKALQIPV
ETLSQAQSPSPSNSTDPAPVQTYQVYGLHATGSGKSTKIPVAAVYGVNVLVLP
SVAAITGSGPMSRAYTDPNRTNGTNTVITGAKLYSTYTGKFLADGGSGGAYDVLI
CDECHAQATSTILGTVLDQAEATVGLTVLATATPGSITVPHSNIEEVLVSGITG
IFPYKATP IALLKGRHILIFCHSKKICDEIAGRLGMLNVAARYGLDVSIFPTG
DIVVACATAMTGVTFGDFDSVILDCNAVEQVDFSLDPTFSEITRPAQDAVRSQR
TRGRGRGLTRYVAGPERSGMSDFVLCYDCAGSWYDIQPAETTVRLRAYLSTP
GLVPQDHLFWESVFTGLTHIDAEFLSOTKQOGLNFSYLTAAQTVCARAQAPPSV
DEMVKLRLKPTLHGLGCVVIVGHIELGGKPALVPDEKLVLYQQYDMEECQAAPYIEQ
LGCVTLAALAYCLSVSGCVVI VGHIELGGKPALVPDEKLVLYQQYDMEECQAAPYIEQ
AQAIAHQFEKVLGLQATQOAVI BEI VATNNQKLEAFWHHGHMNFVSGLYQLAGL
PTLPGNPAVLMMLATGASVTSPLTNTQTMFENILGGWVATHLAGPOSSFAVVGSLAG
GVTGAILGRVLLDILAGVAGSALVAFKIMGELPTTEDLVNLLPAILSPGALV
GUTCAIIRGVHGPCEGAVOMNRLIAPASGRNHVSPTHVPSDDAAKVTALLSLSL
VTSLLRHLQWINEBYPSCDDWILRIIWDKWCVSADFKTWSAKIMPALPGLPFIIS
CQGYGVWVRGQGVITTRPCGATITIVDKWCSVLADFKTWSAKIMPALPGLPFIIS
PSTPCPSNYTALMARVANSVYEVRRVGDTHYITGATEDBELKCPQVPAAREFTVE
GVRLRYAPPCPLPDLITFMVGLNSYAIQSOLPCEPDEPVSIVLSMLRQENSHIT
TAARRLARGSPLEASASQASLPKATQOTHRPHDAELVDANILLMRQENSHIT
RVSESTVILVDSFPLRAQTDAAELVAAECFKPKPKYPALPIWARPDYNNPLDLR
WKTYPVVPVTPALPRGAPVPPPRKRTIOLDGNSVSAALAAAKKSPFSPVNPQ
DENSSGVDYTSSTSKVPPSPGESDESSESSMPLEGEPPDPLSCDSWSTVSDS
EEOGVSCSWSYTGALITPCSAEEXKLPI SPLNSLIRHNLVYTSRSASQXQK
KVTFDLRLQDLDDHITKAEVRAKMLIETIACALVPHSARSKFYSAKDV
RSLSSRAIDQIYRSEWDLLEDTPITTIIMAKNEVFCVDPARGGRKRBALVYVPLG
VRVCEKRVLDIYQKLSITMGTAIFYQFYSQVPERILLKMTWSTKTPGLGSDYTRCF
GLVTFQEDRVEBEIYQCCNLPEARKVISLITELRYCGCFMFSKGAQGYRRRARS
DSVTPTEGNTITCVI KACAAKAAALRNPDFLVCGDILVVAASDGVDEDTALRAAE
EAMTRYSAPEGDAPOQYTDLELITCSSNVSVALDNKGKYYILTRDATTPLARAAME
TARHTPSWHLGNIMYKATIMVWMMTHFPESILQSOEILDRPLDEMYGATVYTP
LDLUPAIIRLHUGLUSATFLHSGVPELVNRVAGTFLKLGCPPLRAMHRBARAVRAKLIAQ
GGRAKICGLYLFNVAVRTKTNITLPLPATQGLQDLSWFTVGVGGNDIYHVSIRARTRYVL
LLCLLLLTGVGIFLLPAR"

mat_peptide	299 . . 871
	/product="core protein"
mat_peptide	872 . . 1447
	/product="envelope protein"
mat_peptide	1448 . . 2503
	/product="Ns1 protein"
mat_peptide	2504 . . 3334
	/product="Ns2 protein"
mat_peptide	3335 . 5287
	/product="Ns3 protein"
mat_peptide	5288 . . 5449
	/product="Ns4a protein"
mat_peptide	5450 . . 6232
	/product="Ns4b protein"
mat_peptide	6233 . . 7588
	/product="Ns5a protein"
mat_peptide	7589 . . 9361
	/product="unssh protein"

[illegible]

QY	62	TCGTCGGCGCTCCCGTAGGAGCGTGCAGAGCCCTTCGCGAATGCGGTGAGGGCCCTTG	121
Db	714	TCGTCGGCGCTCCCGTAGGAGCGTGCAGAGCCCTTCGCGAATGCGGTGAGGGCCCTTG	773
QY	122	AAGACGGGATAAAATTCGCGAACHAGGAATTTGCCCGTTGCTCCTTTCTATTTTCTCTC	181
Db	774	AGGACGGGATAAAATTCGCGAACHAGGAATTTGCCCGTTGCTCCTTTCTATTTTCTCTC	833
QY	182	TGCTCTGTTCTTCTTGCTTAAATTCATCCAGCAGCTAGTCTAGAGTGCGCGAATACGTTG	241
Db	834	TTGCTCTATTCTCTTGCTTAAATTCATCCAGCAGCTAGTCTAGAGTGCGCGAATACGTTG	893
QY	242	GCCTCTATGCTCTTACCAACGACGTTTCCAAATAGCAGTATTTGTAGCAGCGCGATGACG	301
Db	894	GCCTCTATGCTCTTACCAACGACGTTTCCAAATAGCAGTATTTGTAGCAGCGCGATGACG	953
QY	302	TTATTTGCGACACACCGGCTGCATCCTTTGTGTCAGGACGGCAATACATCCACGTTGCT	361
Db	954	TCATTTGCGACACACCGGCTGTGTACTTTGTTTCAAGACGGCAATACATCCACGTTGCT	1013
QY	362	GGACCCAGTCACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACGCTTCGATAC	421
Db	1014	GGACCCAGTCACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACGCTTCGATAC	1073
QY	422	GCAGTCATGTGGACCTATTAGTGGCGGCGCCAGATGCTCTCGCTCTACGTTGGGTG	481
Db	1074	GCAGTCATGTGGACCTATTAGTGGCGGCGCCAGATGCTCTCGCTCTACGTTGGGTG	1133
QY	482	ACATGTGTGGGCTGTCTTCTCTCGTGGACAAAGCTTCACGTTTCAGACCTCGTGGCCATC	541
Db	1134	ATATGTGGGAGCGCTCTTTCTCGTGGACAAAGCTTCACGTTTCAGACCTCGTGGCCATC	1193
QY	542	AAACCGGTCAGACCTGTAACTGCTCGCTGTATCCAGGCGCATCTTCAGGACATCGAATGG	601
Db	1194	AAACCGGTCAGACCTGTAACTGCTCGCTGTATCCAGGCGCATCTTCAGGACATCGAATGG	1255
QY	602	CTTGGGATATGATGAACTGGT	625
Db	1254	CTTGGGATATGATGAACTGGT	1277
RESULT 8			
HPREGS			
LOCUS	Hepatitis C virus complete genome sequence.		
DEFINITION	D17763		
ACCESSION	D17763.1		
VERSION	GI:514395		
KEYWORDS	polyprotein; C; E1; E2/NS1; NS2; NS3; NS4a; NS4b; NS5a; NS5b.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae		
AUTHORS	Hepacivirus.		
TITLE	1 (bases 1 to 48)		
JOURNAL	Sakamoto, M., Akahane, Y., Tsuda, F., Tanaka, T., Woodfield, D.G. and		
MEDLINE	Sakamoto, H.		
PUBMED	Entire nucleotide sequence and characterization of a hepatitis C		
REFERENCE	virus of genotype V/3a		
AUTHORS	J. Gen. Virol. 75 (Pt 7), 1761-1768 (1994)		
JOURNAL	8021605		
REFERENCE	2 (sites)		
AUTHORS	Sakamoto, M.		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 9456)		
AUTHORS	Okamoto, H.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-SEP-1993) Hiroaki Okamoto, Jichi Medical School,		
REFERENCE	Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi		
AUTHORS	329-04, Japan (E-mail:hokamoto@jichi.ac.jp).		
JOURNAL	Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)		
FEATURES	Location/Qualifiers		
SOURCE	1..9456		
	Hepatitis C virus"		

/mol_type="genomic RNA"		/product="NS4a"		mat_peptide	
/isolate="NZL1"		5491..6273		5491..6273	
/db_xref="taxon:11103"		/product="NS4b"		/product="NS4b"	
1..339		6274..7629		6274..7629	
340..9405		/product="NS5a"		/product="NS5a"	
/codon_start=1		7630..9402		7630..9402	
/product="polyprotein"		/product="NS5b"		/product="NS5b"	
/protein_id="BAA04609.1"		9406..9425		9406..9425	
/db_xref="GI:1183029"		polyA_site		polyA_site	
/translation="MSTLPKPKQKTKRNTIRPQDKVPGGQGVGVVYVLRPRRL GVRATRSRQPRQPIKARRSEGRSQAQVPMPLVYNGSCGAGWLLSPRG SRPSWGNDRPRRNLGKVIITLTCGPADLMGYIPLVGPVGVARALAGVRALED GINFATGNLPGCSFSIFLALFSLIHPAASLEWNTSLGLVLTNDSCNSSIYEADD VILHTPGVPCVQDQNTSTCTWPTVAVRTVGAITASIRSHVLLVGAATWASALY VDMCGAVTQDQFTFRRHQTQVOTNCSLYPGLSHGRMAWMMNWSVAVGVV AHVLRPLFDIMAGAHGILAGLAYSQMNWAKVAIIMVFSFGDAHTYTTGGTA SRHTQAFAGLFTGQQKQLVNTGSHWINSITNLCNSESINTGFIAGLIFYHKFNST GCPQLSSCKPTTFRQMGWPLTDANIITGSDRPYCWHYAPRCDIIVPASSVCGPVY CTPSPVVVGTTDARGVPTYTWGENEKOVFLKSKQRPSPGRFSGSWNNTSFLTKCG APPONIYGEENPHNESDLFCPTDFRKHPTTYSRGCGPMLTPRCMVNDYPIRLWHY PCTVDFRLKVMFVGGFHEHRTAACNMTWGRGCRDIEDRDRSEQHLLSTTELAILP CSFTMPALSTGLIHLHQNIVDQYLYGVSGMGWALKWEFVILFVLLADARVCVA LWMLMSOTEALLENVLTNAVAAAGTHIGWLVAFCAWYVKGKLPVLTYSLTG LWSLALLVLLPQRAYANGSDSATLGGVILVFGFETLSPWYKMGRLMWNYOTI CEESALHWVPPPLARGSRDGVILLTSLYPSLIEDITKLLMAVLGPLYLIQAITT TPYVRAHVLRCLMVKLVSGKGYFQMIILSIGRWENTLYDHLAPQHWAAAGLKD LAVATEPVLFSMEIKVITWAGDTAAACGDLICGLPVSARLREGLVLPADDYREMGWR LLAPITAYAAQTRGLGTVISLTGRDKNVVTGEVQVLSATQTFGLTGVGVITWTV HAGSRTLAGAKHPALQMYTNDVQDLVGPAPPGALESLEPCACGSDLYLVTRDADVI PARRGSDTASILSRLPCLAKSGSGPVMPGSHVAGI FRAAVCTRGVAKSLQFIPV ETLSQARSPPSNDSTPAVQSQVGYLHAPTSGSKTKVPAAYVAGQYVNLVNP SVATLGFSGFMSRATGIDPNRGTNTVTTAKUTYSTYKGLADGCGSGAYDVI CDECAQDATSLIGTLDQAEATAGVRLTATATPGSITVPSHNTIEEVALGSEGE IPFYGAIPALLGGRHLIFCHSKKCCDEIASKLGRMGLNNAVAYRGLDVSIVPTTG DVVGTALMTGQDFDVIDCNVAVEYVDFLDPFTSEITETAPQDAVRSORR GRTGRGLGTYRYVAGSRRPSCGFVSLCEYDAGCSWYDLOPATVRLRAYLSTP GLPVCQDHLDFWESVFTGLTHIDAFHSQTKQGLNFSVLTAYQVNCARQAQPPSW DEMKCLVRLKPTLHGTPPLLYRLGVPQNETCLTHPTIKYLMACMSADLEVTSTWVL LGGVLAALAAVCLSGVCVVI VGHIELEKPAVLPDKEVLYQYDDEMSQAPYIEQ AQVIAHQFKEILMLQATQOQAVIEPTVTTNWQLEAFWFKHMMNFVSGIYIYLAGL STLPNPAVPAWMAFASVTSPLTNTQMFNILGWVATHLAGPQSSAFVVSGLAG AAIGIGLGRVLLDILAGVAGVSGALVAFKIMGCECTPAEDMWNLLPALLSPGALV GVICAILRHVGVGEGAVQWNRLLIAPASRGNHVSPTHYPESDAARVATLLSSLT VTSLLRLHQWLNEDYPPSCDDMLRTIWDVCSVLADPFKWLAKIMPALPGLPFIIS COKGYKVRGWDGSTRPCGAAITGHVKGSMRLAGPRTCANMWHGTFPTNEYTTG PSTPCSPNYTRALNRWANSYVEVRVGDHYITGATEDELKCPQVPADEFTEVD GVRLHRYAPCPKPLARDITFMVGLHSYITIGSOLPCEPEPDVSVLTSLMRDPSHITAE TAARLARGSPSEASSASQLSAPLSKATCQTHRPHDAELVDANLLWRQBMGNIT RVSEETKVVVLSFPLRATDDVFSVAECFKPPKTPPALPIIWARPDYNNPDLDR WKAPDYVPTVHGCLAPRGAPPVPPRRKRTIQDGNVNSAALAAEKSPPSKPQ EENSQSGVDYDTSSTSKVPPSGSESDSCSMPLEGEPEPDLSGDSWSTVSDS ERQSVCCSMYSWTGALITPCSABEELPI9PLNSLLRHHNLVYSTSRASQROK KYTFRLQVLDHYKTKAVERASRVKARMLTIEEACALVPHSARSKYSADV RSLSRANQIRSVWEDLLEDTTPTPTIMAKNEVFCVDPKAGRKPRLIIVYEDLG VVCSEKRALDIYQKLSIETMGPAFYQFSPOORVERLLKMTWTKPTFGFYDTRCF DSTVTEQDIRVEIKLSYQCNLEPEARKYISSILTERLYCGGFNFSGKQCYRRCAR GYLPSFGNTICTYKATAAARALNRNPDVLCVGDVVAESGVEDDRAALRAFT EAMTRYSAPGPAPQATYDLELITSCSNVSVARDDKGRYYLTRDATTPLARAWE TARHTPVNSWLNIIIMYARTINRVMWMTFFSILQSEILDRPLDFEMYGATYSVTP LDLPAIERLHGLSALTHLSYFVELNRVAGTLRKGLCPPLRAWRHRARAVRAKLIAQ GKAKICGLYLFNFAVRTKNTLPLPAAGQLDLSWFTVGVGNDIYHSVSRARTRLH LILCLLLTVGVGIFLLPAR"		BASE COUNT			
1998 a		2684 c		2578 g	
2196 t		ORIGIN			
Query Match.		93.0%; Score 585.6; DB 14; Length 9456;			
Best Local Similarity		96.2%; Pred. No. 1.4e-152;			
Matches 600; Conservative 0; Mismatches 24; Indels 0; Gaps 0;					
QY	2	TCGGTAAGGTCTATCATACCCCTTACGTGGGATTCGCGGATCTCATGGGGTACATCCCGC	61		
DB	695	TGGGTAAAGTCAATCATACCCCTTAACTGCGGATTTGCCGACCTCATGGGGTACATCCCGC	754		
QY	62	TCGTCCGGCTCCCGTAGGAGCGTTCGCAAGAGCCCTTTCGCGATGCGGTAGGGGCCCTTG	121		
DB	755	TCGTCCGGCTCCCTGTAGGAGCGTTCGCAAGAGCCCTTCGCGATGCGGTAGGGGCCCTTG	814		
QY	122	AAGACGGGATAAATTCGCAACAGGAAATTCGCCGGTTCCTTTCTATTTCTCTTC	181		
DB	815	AAGACGGGATAAATTCGCAACAGGAAATTCGCCGGTTCCTTTCTATTTCTCTTC	874		
QY	182	TCGCTCTGTTCTCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGGAATACGTCTG	241		
DB	875	TTGCTCTGTTCTCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGGAATACGTCTG	934		
QY	242	GCCTCTATGTCCTTACCACGACTGTTCCAAATAGCAGTATTGTGTACGAGCGCGATGAGC	301		
DB	935	GCCTCTATGTCCTTACCACGACTGTTCCAAATAGCAGTATTGTGTACGAGCGCGATGAGC	994		
QY	302	TTATTCTGCACACACCCGGCTGCATACCTTGTGTCCAGACGGCAATACATCACTGCT	361		
DB	995	TCATTTCTGCACACACCCGGCTGTGTACCTTGTGTCCAGACGGCAATACATCACTGCT	1054		
QY	362	GGACCCCACTGACCTACAGTGGGAGTCACTAGTCCGAGCAACACCCGCTTCGATAC	421		
DB	1055	GGACCCCACTGACCTACAGTGGGAGTCACTAGTCCGAGCAACACCCGCTTCGATAC	1114		
QY	422	GCAGTCATGTGGACCTATTAGTGGGCGCGCCACCATGTGCTCTGCGCTCTACGTGGGTG	481		
DB	1115	GCAGTCATGTGGACCTATTAGTGGGCGCGCCACCATGTGCTCTGCGCTCTACGTGGGTG	1174		
QY	482	ACATGTGTGGGCTGTCTTCTCGTGGGACAAAGCCCTTACAGTTCAGACCTCGTGGCCATC	541		
DB	1175	ATATGTGTGGGCTGTCTTCTCGTGGGACAAAGCCCTTACAGTTCAGACCTCGACGCCATC	1234		
QY	542	AAACGGTCCAGACCTGTAACCTGCTGCTGACCCAGGCCATCTTTTCAGGACATCGAATGG	601		
DB	1235	AAACGGTCCAGACCTGTAACCTGCTGCTGACCCAGGCCATCTTTTCAGGACATCGAATGG	1294		
QY	602	CTTGGGATATGATGATGAATCGT	625		
DB	1295	CTTGGGATATGATGATGAATTTGT	1318		
RESULT 9		HPCUS114CE		1786 bp	
LOCUS		Hepatitis C virus genes for C, E and E2/NS1, partial cds.		RNA	
DEFINITION		Hepatitis C virus genes for C, E and E2/NS1, partial cds.		linear	
ACCESSION		D14309		VRL 01-FEB-2000	
VERSION		D14309.1			
KEYWORDS		C; E; E2/NS1; core protein; envelope protein; nonstructural protein.			
SOURCE		Hepatitis C virus			
ORGANISM		Hepatitis C virus			
		Viruses; serNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			

5'UTR
CDS

mat_peptide
mat_peptide
mat_peptide
mat_peptide
mat_peptide
mat_peptide


```

REFERENCE
AUTHORS      Okamoto,H., Tokita,H., Sakamoto,M., Horikita,M., Kojima,M.,
              Iizuka,H. and Mishiro,S.
TITLE        Characterization of the genomic sequence of type V (or 3a)
              Hepatitis C virus isolates and PCR primers for specific detection
JOURNAL      J. Gen. Virol. 74 (Pt 11), 2385-2390 (1993)
MEDLINE      94065664
PUBMED       7504073
REFERENCE    2 (bases 1 to 1786)
AUTHORS      Okamoto,H.
JOURNAL      Submitted (28-JAN-1993) to DDBJ by:
COMMENT      Hiroaki Okamoto
              Immunology Division
              Jichi Medical School
              Kawachi
              Tochigi 329-04
              Japan
              Phone: 0285-44-2111x3334
              Fax: 0285-44-1557.
              Location/Qualifiers
                1. .1786
                  /organism="Hepatitis C virus"
                  /mol_type="genomic RNA"
                  /isolate="US114"
                  /db_xref="taxon:11103"
                  <1..277
                    /citation=[1]
                    278..>1786
                      /codon_start=1
                      /product="polyprotein precursor"
                      /protein_id="BAA03255.1"
                      /db_xref="GI:471113"
                      /translation="MSTLPKPKQKTKRNTIRPQDKVFGGQIVGGVYVLPRRGPRL
                      GVPATKTSERSOPRRGPIPKARSEGRSQAQGYPMPLYNEGCGWAGWLLSPRG
                      SRSWGNDRRRSRNLGVLDITCGFADLMGYIPLVGAPVGGVARLAHGVRALED
                      GINFATGPGGCSFSLIALSLHLPASLEWNTSGLYVLTNDSCNSIIVYBAD
                      VILHTPGCIPCVQDGNVSRCTVPTVAVRVGATASIRSHVDLLVGAATWCASALY
                      VGMCGAVFLVGAQTFRRRHQTVQCNCLYFPHLSGHRMADMMWNSPAGVMV
                      AHVLRIPQTLFDIVAGAHGILAGLAYISMQGNWAKVALIMVFGSDVASTHTGSM
                      ARDVYRTGFFARGPSQNLQLVNTGSHWIRNLTALNCNESINTGFIAGLLYNNKFNST
                      GCQRLSCKPITSFRQWGSLSLTANITDSSDDKPYCWHYAPRSCDTPASSV"
              mat_peptide 278..850
              /product="C"
              mat_peptide 851..1426
              /product="E"
              mat_peptide 1427..>1786
              /product="E2/NS1"
              BASE COUNT 353 a 522 c 507 g 404 t
              ORIGIN
                Query Match 92.4%; Score 582.4; DB 14; Length 1786;
                Best Local Similarity 95.8%; Pred. No. 9.7e-152;
                Matches 598; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
              QY 2 TGGGTAAGTCATCGATACCTTACGTGCGGATTTCGCGATCTCATGGGGTACATCCCGC 61
              DB 633 TGGGTAAGTCATCGATACCTTACGTGCGGATTTCGCGGACCTCATGGGGTACATCCCGC 692
              QY 62 TCGTCGGCGCTCCCGTAGGAGCGGTTCGCAAGAGCCCTTGGCGATGGCGTAGGGCCCTTG 121
              DB 693 TCGTCGGCGCTCCCGTAGGAGCGGTTCGCAAGAGCCCTTGGCGATGGCGTAGGGCCCTTG 752
              QY 122 AAGACGGGATAAATTCGCAACAGGGAATTCGCGGTTGCTCTTTTCTATTTCCTTC 181
              DB 753 AAGACGGGATAAATTCGCAACAGGGAATTCGCGGTTGCTCTTTTCTATTTCCTTC 812
              QY 182 TCGCTCTGTTCTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGCTG 241
              DB 813 TTGCTCTGCTCTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGCTG 872
              QY 242 GCCTCTATGCTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACAGAGCCGATGACG 301

```

```

Db 873 GCCTCTATGCTCTTACCAACGACTGTTCCAAATAGCAGCATTTGTATGAGGCCGATGACG 932
QY 302 TTATTCTGCACACACCCGGCTGCATACCTTGTGTCCAGACGGCAATACATCCAGTGTCT 361
Db 933 TTATTCTGCACACACCCGGCTGTATACCTTGTGTTCAGACGGCAATGTATCCAGTGTCT 992
QY 362 GGACCCCACTGACACCTACAGTGGCAGTCAAGTACGTTCGGAGCAACACCCGCTTCGATAC 421
Db 993 GGACCCCACTGACACCTACAGTGGCAGTCAAGTACGTTCGGAGCAACACCCGCTTCGATAC 1052
QY 422 GCAGTCATGTGGACCTTATTAGTGGGGCGGCGCACAGTGTGCTCTGGCTCTACGTGGGTG 481
Db 1053 GCAGTCATGTGGACCTTACTAGTGGCGCGCCACGATGTGCTCTGGCTCTACGTGGGGG 1112
QY 482 ACATGTGTGGGCTGTCTTCTCGTGGGACAGCCCTTACAGTTCAGACCTCGTCGCCATC 541
Db 1113 ATATGTGTGGGCGGCTCTTCTCGTGGGACAGCCCTTACAGTTCAGACCTCGTCGCCATC 1172
QY 542 AAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCGCATCTTTTCAGGACATCGAATGG 601
Db 1173 AAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCGCATCTTTTCAGGACATCGAATGG 1232
QY 602 CTTGGGATATGATGATGAATGGT 525
Db 1233 CTTGGGATATGATGATGAATGGT 1256

RESULT 10
HPCTH85CE 1786 bp RNA linear VRL 01-FEB-2000
LOCUS      Hepatitis C virus genes for C, E and E2/NS1, partial cds.
DEFINITION D14307
ACCESSION  D14307.1 GI:456468
VERSION    C; E; E2/NS1; core protein; envelope protein; nonstructural
KEYWORDS   protein.
SOURCE      Hepatitis C virus
ORGANISM   Hepatitis C virus
REFERENCE   1 (sites)
AUTHORS     Okamoto,H., Tokita,H., Sakamoto,M., Horikita,M., Kojima,M.,
              Iizuka,H. and Mishiro,S.
TITLE       Characterization of the genomic sequence of type V (or 3a)
              Hepatitis C virus isolates and PCR primers for specific detection
JOURNAL     J. Gen. Virol. 74 (Pt 11), 2385-2390 (1993)
MEDLINE     94065664
PUBMED      7504073
REFERENCE   2 (bases 1 to 1786)
AUTHORS     Okamoto,H.
JOURNAL     Submitted (28-JAN-1993) to DDBJ by:
COMMENT     Hiroaki Okamoto
              Immunology Division
              Jichi Medical School
              Kawachi
              Tochigi 329-04
              Japan
              Phone: 0285-44-2111x3334
              Fax: 0285-44-1557.
              Location/Qualifiers
                1. .1786
                  /organism="Hepatitis C virus"
                  /mol_type="genomic RNA"
                  /isolate="Th85"
                  /db_xref="taxon:11103"
                  <1..277
                    /citation=[1]
                    278..>1786
                      /codon_start=1
                      /product="polyprotein precursor"
                      /protein_id="BAA03253.1"
                      /db_xref="GI:471118"
                      /translation="MSTLPKPKQKTKRNTIRPQDKVFGGQIVGGVYVLPRRGPRL

```


GVRATKTSERSQPRGRQPIPKARRSRSWAQPCYPWPLYGNEGCGWAGLLSPRG SRPSGPNDRPARRSRNLGKVIDTLCGFADLMGYIPLVGAPVGGVARALAGHVRALD GINATFNLCPCSPSIFILLALFSLIPAALEWENTSGLYLTNDSCNSIVCBAD VILHTPGCPCVHNGNTCTWTPTVAVRHVGATTASIRSHVDLLVGAATMCSALY VGDCGAVFLVQAFTFGRHOTVQTCNCSLYPGHLSGHRMAMMMNWSPAVGMV AHLVRLPOTLFDIIAGAHWGLIAGLAYYSMOGNWAKVAIIMVPSGDAETRTTGGTA AHDVYRPSLFRNARONLQLVNTNGSHINRTALNCNDSINTGFIARLHYHFNST GCPQLSSCKPTPSNQSGPLTDBNITGPDSDDKPYCHYAPRCKSVPASSV" 278...850 /product="C" mat_peptide 376 a 526 c 490 g 394 t 851...1426 /product="E" mat_peptide 376 a 526 c 490 g 394 t 1427...>1786 /product="E2/NS1" BASE COUNT 376 a 526 c 490 g 394 t ORIGIN									
Query Match 91.4%; Score 576; DB 14; Length 1786; Best Local Similarity 95.2%; Pred. No. 6e-150; Matches 594; Conservative 0; Mismatches 30; Indels 0; Gaps 0;									
Qy	2	TGGGTAAGGTATCATGATACCCCTTACGTGGGATTCGCCGATCTCATGGGGTACATCCCGC	61						
Db	633	TGGGTAAGGTATCATGATACCCCTTACGTGGGATTCGCCGATCTCATGGGGTACATCCCGC	692						
Qy	62	TCGTGGGCTCCCGTAGAGCGCTGCGAAGAGCCCTTGGCGATGGCGTGAGGGCCCTTG	121						
Db	693	TCGTGGGCTCCCGTAGAGCGCTGCGAAGAGCCCTTGGCGATGGCGTGAGGGCCCTTG	752						
Qy	122	AAGACGGGATAAATTCGCAACAGGGAATTCGCCGGTGTGCTCCCTTCTATTTTCCCTC	181						
Db	753	AAGACGGGATAAATTCGCAACAGGGAATTCGCCGGTGTGCTCCCTTCTATTTTCCCTC	812						
Qy	182	TCGCTCTGCTCTCTTCTGCTTAATTCATCCAGCAGCTAGCTAGAGTGGCGGAATACGTCG	241						
Db	813	TTGCTCTGCTCTCTGCTTAATTCATCCAGCAGCTAGCTAGAGTGGCGGAATACGTCG	872						
Qy	242	GCCTCTATGCTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAG	301						
Db	873	GCCTCTATGCTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAG	932						
Qy	302	TTATTTCTGCACACACCCCGCTCATACCTTGTGTCAGGACGGCAATACATCCAGTGT	361						
Db	933	TTATTTCTGCACACACCCCGCTCATACCTTGTGTCAGGACGGCAATACATCCAGTGT	992						
Qy	362	GGACCCAGTGCACACTACAGTGGCAGTCAAGTACGTCCGAGCAACACCGCTTCGATAC	421						
Db	993	GGACCCAGTGCACACTACAGTGGCAGTCAAGTACGTCCGAGCAACACCGCTTCGATAC	1052						
Qy	422	GCAGTCATGTGACCTATTAGTGGCGCGGCCACAGTGTCTCGCTCTACGTGGGTG	481						
Db	1053	GCAGTCATGTGACCTATTAGTGGCGCGGCCACAGTGTCTCGCTCTACGTGGGTG	1112						
Qy	482	ACATGTGGGGCTGTCTTCTCGTGGGACAGCCTTACGTTACAGCTCTCGCCATC	541						
Db	1113	ATGTGTGGGGCTGTCTTCTCGTGGGACAGCCTTACGTTACAGCTCTCGCCATC	1172						
Qy	542	AAACGGTCCAGACCTGTAACCTGCTGCTGTAACCCAGGCCATCTTTCAGACATCGAATGG	601						
Db	1173	AAACAGTCCAGACCTGTAACCTGCTGCTGTAACCCAGGCCATCTTTCAGACATCGAATGG	1232						
Qy	602	CTTGGGATATGATGATGAATGGT	625						
Db	1233	CTTGGGATATGATGATGAATGGT	1256						
RESULT 11 AF046866 LOCUS DEFINITION ACCESSION VERSION KEYWORDS									
AF046866 Hepatitis C virus type 3a (CB), complete genome. 9425 bp RNA linear VRL 19-FEB-1998 AF046866 AF046866.1 GI:2895898									

SOURCE ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1 (bases 1 to 9425)
Shukla,D.D., Chaturvedi,S., Cao,J.Y. and Hoynes,P.A.
Complete Nucleotide Sequence of the genome of Hepatitis C Virus
type 3a (CB)
Unpublished
2 (bases 1 to 9425)
Shukla,D.D., Chaturvedi,S., Cao,J.Y. and Hoynes,P.A.
Direct Submission
TITLE
Submitted (04-FEB-1998) Biomolecular Research Institute, 343, Royal
Parade, Parkville, Melbourne, Victoria 3052, Australia
LOCATION/Qualifiers
1..9425
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="type 3a (CB)"
/db_xref="taxon:11103"
340...9405
/codon_start=1
/product="polyprotein"
/protein_id="AA03058.1"
/db_xref="GI:2895898"

CDS

/translation="MSTLPKPKTKRNTIRPDQVPPGGQIVGVVYVLPRRGRRL
GVCATKTSERSQPRGRQPIPKARRSRSWAQPCYPWPLYGNEGCGWAGLLSPRG
SRPSGPNDRPARRSRNLGKVIDTLCGFADLMGYIPLVGAPVGGVARALAGHVRALD
GINATFNLCPCSPSIFILLALFSLIPAALEWENTSGLYLTNDSCNSIVCBAD
VILHTPGCPCVHNGNTCTWTPTVAVRHVGATTASIRSHVDLLVGAATMCSALY
VGDCGAVFLVQAFTFGRHOTVQTCNCSLYPGHLSGHRMAMMMNWSPAVGMV
AHLVRLPOTLFDIIAGAHWGLIAGLAYYSMOGNWAKVAIIMVPSGDAETRTTGGTA
AHDVYRPSLFRNARONLQLVNTNGSHINRTALNCNDSINTGFIARLHYHFNST
GCPQLSSCKPTPSNQSGPLTDBNITGPDSDDKPYCHYAPRCKSVPASSV"
CTPSPVVGTTDAKVPYTMGANDTDFELLESIRPGGRWFGCTMNSTFVKTCG
ASPCDIYGGGNSGNEBILFCTDCFKHPEATYRCGAGPWLTPRCWDYVRLWHY
PCTVNTLTKVRMVGSEHRTAACNWRGERCDIEDRSEQHPLHSTTELAALP
CSFTMPALSTGLIHLQNVIVDQYLYGSGMVGWALKMEFVILI FLLADRRVCVA
LWMLMTQAABALHNLVNAVAAGTHGIGWYLVAFCAAWYRGKLVVTVYSLTG
LWPLAILVLPORAYAWSGDSATLGAGVLALFGFTLSLPMWIKGLIILWNOVTI
CRESALOVVPPILLARGSRDQVLLTSLYPSLIFDITKLIIVGLYLILQAQVTT
TPYFVRAQVLRCLMVRSMVGKGYFMILISIGRNFNTLYDHLAPWQWAAAGLKD
LAVATEPIFSPMEIKVITWGADTAACDILGGLPVSARLGEVLLGPADDYREMGWR
LLAPATVIAQQTGGLITVSLTGRDNVTGEVQLSTATQTFTLGTVGVTWTV
HGASRTLAGAGHPALQWNTVDQLWGMPAPPAGAKSLPCACGASDLYLVTRDADVI
PARREDSTASLLSPRLACLKSGSGGPMCPSGHVAGIFRAAVCTRGAKALQFIPV
ETLSQTSPSPSDNSTPAPVQSYQVGYLHAPTSGSKSTKVPAYAAQYTVVLNP
SVATLGFSGFMSRAYGIDPNIRNVTGAKLTISTYKFLADGGCGGAYDII
CDEHAQDATSLIGITVLDQATAGVRLTVLATTPGSIIVPHSNIEEVALGSEGE
IPFYKATPIALLKGRHLIFCHSKKCDDEIASKLRGMGLNAYAYRGLDVSVIPTTG
DVVVCATALMTGTGDFSDVIDCNVAQVYDFSLDPTFSETTAPDADVSRSORR
GRTGRGLGTVRYVAPGRPSGMPDSVILCEYDAGCSWYDLOQAEVTVRLRAYLSTP
GLPVCQDHLDFWESVFTGLTHDAHPLSQTQKQGLNFSLYLTAYQATVCARAQPPSW
DETWCLVRLKPTLHGPTPLLYRLGFPVQNEICLSHPTIKYNMACSDALEVTTSTWVL
LGGVLAHAALVCLSVGCVVIIGHIELGKPALVPDKVLYLQYDMEESQSAPIEQ
AQATLAHQFEKVLGLQRTAQQAIEPIVATNMOKLEAFWKKHMMVFSGQYLAEL
STLPCNPAVASLMAFTASVTSPLTNTMFFNIGWVATHLAGPQSSAFVSVGLAG
RAIGIGIGRVLDDILAGYAGVSGALVAFKIMGSELPTAEDMNLPAILLSPGALV
GYICAIIDRRHVPGEGAWMNRLLIAFASRNHVSPTHYVPESDAAARVTALESILT
VTSLLRLHUNEDYDPSGSDNWLRIIWDWVCVSLSDPSKWSLAKIMHGLPGLPFIS
COKYQSPNYTRALMVAANSYVEVRVGDPHYITGATEDELKPCQVAPAEFPFTEVD
SVTLHRYPACKPLAREBITFVSGLNSVTIGSQLCEPEPDVSLTMLRDPSSHITAE
TAARLARGSPSEASSASQLSASLKATCQTHRPDPAELVNLNLRWQMSNIT
VSESTKVILDSPEFLRAETDDELSVAACECFKKPKYPPALPIWAPDYNPPLDOR
WKADYVPVTHGHCALPRGAPPVPPRRKRTIQDGSNVSAALRAKSPSLKQP
EENSSGVDTSQSTTSKVPSPGSESDSCSNMPLGEPEGDPDSCDSSTVSQSDS
EEQSVVCSMSYMTGALITPCSAEEKLPISPLNSLRLHNLVYSTSASORQR
KVTDFRLOVLDHDKYTLKVEKASRVKARMLTIEACALVPHSARSKFCYSKDV
RSLSSRAINQIRSWEDLLEDTTPTTIIMAKNEVFCVDPKAGKRPKARLIVYDFLG
VRVCEKRALYDVIQKLALETIGSAFQFQSPQORVERLLMKMTSKTTFGFSYTRCF
DSTVTEQDIRVEETIYQCNELEPEARKVISLTERLYCGGPMFNSKGAQGRRCRAS
GVLPTSGFNTTCYIKATAAANGAGLRDPLVLCVGDLLVVAESDGVDEDDGAALRAFT

BASE COUNT		1979 a	2699 c	2599 g	2148 t
ORIGIN		BAMTRYGAPGDAPOPTDYDELELITSCSSNVSVARDDKGRYYVYLTRDATTPLARAWE TARHPVNSWGNLIIMTAPTIWVWMTHEPESLOSQELDRPLDFEMYGATYVTP LDLPAITERLHGLSFTLHSTYSELNARVAGTILKGCPLRWRHRAARAKLIAQ GKAKICGLYFNWAVTKTKTLPFRAGLDLSIFWFTVGNGNDIYHVSARATRYL LCLLLLTGVGIFLLPAR"			
Query Match		91.4%;	Score 576;	DB 14;	Length 9425;
Best Local Similarity		95.2%;	Pred. No. 6.6e-150;		
Matches 594;		Conservative 0;	Mismatches 30;	Indels 0;	Gaps 0;
QY	2	TCGGTAAGTCATCATACCTTACGTGCGGATTCGCGATCTCATGGGGTACATCCCGC	61		
DB	695	TCGGTAAGTCATCATACCTTACGTGCGGATTCGCGACCTCATGGGGTACATCCCGC	754		
QY	62	TCGTGCGGCTCCCGTAGAGCGTCGCAAGAGCCCTTCGCGATGCGGTGAGGGCCCTTG	121		
DB	755	TCGTGCGGCTCCCGTAGAGCGTCGCAAGAGCCCTTCGCGATGCGGTGAGGGCCCTTG	814		
QY	122	AGACGGGATAAATTCGCAACAGGGAATTCGCCGGTGTCTCTTTCTATTTCCTTC	181		
DB	815	AGACGGGATAAATTCGCGACAGGGAATTCGCCGGTGTCTCTTTCTATTTCCTTC	874		
QY	182	TCGCTCTGTTCTTCTCTTAATTCATCCAGCAGTAGTCTAGAGTGGCGGAATACGTCG	241		
DB	875	TTGCTCTGTTCTTCTTAAATTCATCCAGCAGTAGTCTAGAGTGGCGGAATACGTCG	934		
QY	242	GCCTCATGTCCTTACCAACGACTGTTCATAGCAGTATTTGTACGAGCGCGATGAGC	301		
DB	935	GCCTCATGTCCTTACCAACGACTGTTCATAGCAGTATTTGTATGAGCGCGATGAGC	994		
QY	302	TTATTCTGCACACACCGGCTGCATACCTTGTCTCCAGGACGCAATACATCCAGTGT	361		
DB	995	TCATTCTGCACACACCGGCTGGTACTTGTGTTTTCAGAACGACATATATCCAGTGT	1054		
QY	362	GGACCCGAGTGACACTTACAGTGGCAGTCAAGTACGTGCGGACCAACCGCTTCGATAC	421		
DB	1055	GGACCCGAGTGACACTTACAGTGGCAGTCAAGTACGTGCGGACCAACCGCTTCGATAC	1114		
QY	422	GCAGTATGAGCCTATTAGTGGCGCGCCAGCAGTGTCTCTGCGCTCTACGTGGGTG	481		
DB	1115	GTAGTATGAGCCTATTAGTGGCGCGCCAGCAGTGTCTCTGCGCTCTACGTGGGTG	1174		
QY	482	ACATGTGTGGGCTGTCTCTCTCTGTGGGAACAGCTTTCAGACCTCTGCGCATC	541		
DB	1175	ATATGTGTGGGCTGTCTCTCTCTGTGGGAACAGCTTTCAGACCTCTGCGCATC	1234		
QY	542	AAACGGTCCAGCCTGTAACTGCTGTGTACCCAGCGCCATCTTTACAGGACATCGAATGG	601		
DB	1235	AAACAGTCCAGCCTGTAACTGCTGTGTACCCAGCGCCATCTTTACAGGACATCGAATGG	1294		
QY	602	CTTGGGATATGATGATGAATGGT	625		
DB	1295	CTTGGGATATGATGATGAATGGT	1318		
RESULT 12		HPCK3A 9454 bp RNA linear VRL 07-FEB-1999			
HPCK3A		Hepatitis C virus complete genome sequence.			
LOCUS		D28917			
DEFINITION		D28917.1 GI:558520			
ACCESSION		polyprotein.			
VERSION		Hepatitis C virus			
KEYWORDS		Hepatitis C virus			
SOURCE		Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;			
ORGANISM		Hepacivirus.			
REFERENCE		1 (bases 1 to 9454)			
AUTHORS		Yamada,N., Tanihara,K., Mizokami,M., Ohba,K., Takada,A.,			
TITLE		Tsutsumi,M. and Date,T.			
		Full-length sequence of the genome of hepatitis C virus type 3a:			
		comparative study with different genotypes			

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

J. Gen. Virol. 75 (Pt 11), 3279-3284 (1994)
95053917
7964640
2 (bases 1 to 9454)
Date, T.
Direct Submission
Submitted (12-MAR-1994) Takayasu Date, Kanazawa Medical University,
Department of Biochemistry; Uchinada, Kahoku-gun, Ishikawa 920-02,
Japan (Tel:0762-86-2211(ex.3701), Fax:0762-86-4693)
Location/Qualifiers
1. .9454
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="K3a"
/isolate="HCV-K3a/650"
/db_xref="taxon:11103"
340. .9405
/codon_start=1
/product="polyprotein"
/protein_id="BAA06044.1"
/db_xref="GI:558521"
/translation="MSTLPKPKRNTIRRPDVKFPGGGVYVGVVLPKRRGL
GVATRKTSERSQPRRRKPIPKARSEGRSNAQPCVPMPLVGNCEGAGWMLLSPRG
SRFNWPNDRRERNLGKVIDTLCTCGFADLMGYIPLVGAPLGGARALAHVRALED
GINFATGNLPGCSFSIFLLALFSLIHPAASLEWRNTSGLYLTNDCSNSSIVYEADD
VILHTGCI PCVQDNTSTCTPTPTVAVRYGATTASIRSHVDLLVAGAGTWCASLY
VGMTCGPFVLVGOAFTFRPRHRTVQTCNSLYPHLSGQRMWMMWMSPAVGOA
AHILRUPQLFDVAGAHGMIAGLAYSMQGNWAKVAIIMVFSGVDASTHVTAGQA
ARNAYGISLFSVGAQNQLINTNGSMHINTALNCESINTGFTAGLYFHHKFNST
CQQRUSCKPITTFKQGGWPLTDANTGPSDDKPYCWHVAPRCPGIVPALNVCGPV
CFTPVVVGTTDAKGATYTGANKTDFVLLSLRPPSGRWFCTTWNSTGFVLTCLP
APTCNIYGDGRDANESDLFCPTDCFRKHTEATYSCGAGPFWLTPTCLVDIYPRLWHY
PNTFTLFKVRMFVGGHRTAACNWTGRGECEDIEDRDSBQHPHLHSTIELALIP
CSFTPMALSTGLIHLHONIVQVLYGIGSGMGWALKWEFVILIFLLADLACVCA
LMLIITISQAEALENLAVNAAGTGHGWLVAFCAAWYVGRKLVLPVLYSTG
LMSALHVLPLLQRAYAWSGEDSATLGAGILVFGPFTLSPYKHWMLARILWNQYTI
CRCSLHVVPPPLLARGGRDGVILLTSLYPSLI FDI TKLLAALGPLYLIOATTA
TPYFRAHLVRLCMLVRSVMGKGFQMIILSLADGNSNTLYLHDLA PMQWAAAGLKD
LAVATEPVFSPEMEIKVITWGAADTAACDILGLPVSARLGRVLELLGAPADYRMR
LLAPITAYAGVKHQLQWYTNVDOLVGMPPAPGAKSLEPCTCGSADLYLTRADTVI
HAGSRTLAGVKHPLARLKGSSGGPVMCPGSHVAGTSGKSTKPAAYVAQGNVLPV
PARRGSTASLLSPRLARKSGSGPVMCPGSHVAGTSGKSTKPAAYVAQGNVLPV
ETLSQAKSPFSNSPFPAPQSYQVYLHAPHTGSKSTKPAAYVAQGNVLPV
SVAATLPGFSRAYGADPNIRITGNRTVTGAKLTYSTGKFLAGGCGCGADVII
CDDHAODATSLIGITVLDQETAGVRLTVKATKMGKMLGNLNAVAYYRGLDVSVPITG
IPFYKATPIACIKIGRHILFCHSKKCKDKWAKLGRVLELLGAPADYRMR
DVVGCATDMGTGTGSDVSDVNDVAVQVYDFSLDPTFETICTAQDPAVSQR
GRTGRTGTYVYTPGPRSGMDSVVLCEYDAGCSWYDLQATVTRLRAYLSTP
GLPVCODHLDLWESVFTGLTHIDAHFLSOTQAGLNFSLYATQATVCARQAAPP
DSTWCLRLKPTLHPTGLYRGPVONEICLTHPI TKYVWACMSADLEVTITWVL
LGGVLAARVAAVCLSVGCVVI VGHIELGKPLVPDKVLYQYDMEBCSOARPIEQ
AQVIAHQFKEVYLGLOQATQQAIEPIVSNWQKLEVLNKHMMWNSFGIOYLAGL
STLPGNPAVASLMFATSVTPTTNTMFFNLGGWVATHLAGQAASAFVSSLAG
AAIGIGLRLVLLDILAGVAGSALVAFKIMGGEPTTEDMNLPAILSLPGALV
GVIACAILRRHVGPCEGPVQWNRLLIAPASGNHVSAPHYPESDAAARVALLSLLT
VTSLLRLHQNIEDYBSCGDLRLIITWDCVSVSDFKTWSAKIMPALPLPFIIS
CQKYGWVRGDSVMSRCPGASLAGHVKSGMELAGPRTCANMCHGTFPINEYTTG
PSTPCPPNYTRALMRVANSYVRRVGDHYITGATEDGLKCPQVPAFEFTTEVD
GVRIHRYAPCPRLRDEITFMVGLNSVAIGSQLPCEPEPDSVLTSLMRDSSHITAE
TAARLARGPSSEASASQASPLKATCQTHPHDPAELVDANILMRQMGSNIT
RYSESTVILDSFPLRAETDDAELSAACEFKPKPKYPALP IWARDPNYPLLR
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRKMLSTEEACALVPHRSKRSKFGVSAKDV
EQQSVVCCSMYSYGALLTPCSAREEKLPI SPLNSLRLHNLVYTSRSASOROK
ENSSSGVDTQSTASVLPSPGESDESCESSMPLEGECPDLS CDSVSTVSDS
WKSPPVPTVHGCALPPKAPPVPPRRKRIITQDGSNVSAALAAKSPFSKQP
RSLSSKAINQIRSVWEDLEDTTPTTITMAKNRVEFCVDPKAGKKAARILYDPLG
KVTFDLQVLDHHTYKALQEVKERSKVRK

REFERENCE	AUTHORS	JOURNAL	COMMENT
2	(bases 1 to 1786) Okamoto, H. Unpublished Submitted (28-JAN-1993) to DDBJ by: Hiroaki Okamoto Immunology Division Jichi Medical School Kawachi Tochigi Japan Phone: 0285-44-2111x3334 Fax: 0285-44-1557 Location/Qualifiers 1. .1786 /organism="Hepatitis C virus" /mol_type="genomic RNA" /isolate="HEM26" /db_xref="taxon:11103" ..277 /citation=[1] 278. .1786 /codon_start=1 /product="polyprotein precursor" /protein_id="BAA03257.1" /db_xref="GI:471113" /translation="MSTLPKPKTKNTIRRPODIKPGGGIIVGGVYVLRPRRL GVRAATRSERQPRRRQIPKARRSESWAQPGWPLYNCGGAWGLLSRPG SRTPGNDPRRRNRNLGKVIDTLTCTGFADLMGYIPLVGAPVGVARALAHGVRAED GINATPNLPGCSFSI FILLALFSLIHPAASLEWRNTSGLYVLVNDGSSSI VVEAD VILHTPGVPCVQDGNKSTCTWPTVAVRYVATTAIRSHVDLLVGNATTCVSLY VDMGCAVFLVQATFPRRRHQVTCNSCLYPGLSHQMAWMMNWSVPAVGMV AHVRLPOTLEI IAGAHGILAYYSQGNWAKVAVIIVMPSGVDAEYTTTGGS AHATVRSALFTVGARQKLOLVNTNGSHWINSALNCNESINTGFTAGLYPHKFNST GCPHRLSSCKPITSFRQWGLSLTDANISGSSSEDKPYCWHVAPRCTVVPASSV"		
5' UTR			
CDS			
FEATURES	source		
mat_peptide	363 a	517 c	504 g 402 t
mat_peptide			
mat_peptide			
BASE COUNT	363 a	517 c	504 g 402 t
ORIGIN			
Query Match	91.2%;	Score 574.4;	DB 14; Length 1786;
Best Local Similarity	95.0%;	Pred. No. 1.7e-149;	
Matches	593;	Conservative 0;	Mismatches 31; Indels 0; Gaps 0;
QY	2	TGGGTAAGTCAATCATACCTTACGTGCGGATTCGCGGATCTCATGGGTACATCCCGC	61
Db	633	TGGGTAAGTCAATCATACCTTACGTGCGGATTCGCGGATCTCATGGGTACATCCCGC	692
QY	62	TCGTGCGGCTCCCGTAGAGCGCTGCGAAGAGCCCTTGGCATGCGTGGAGGCCCTTG	121
Db	693	TCGTGCGGCTCCCGTAGAGCGCTGCGAAGAGCCCTTGGCATGCGTGGAGGCCCTTG	752
QY	122	AAGACGGGATAAAATTCGCAACAGAGAAATTCGCCGGTTCCTCTTTTCTATTTTCCTTC	181
Db	753	AAGACGGGATAAAATTCGCAACAGAGAAATTCGCCGGTTCCTCTTTTCTATTTTCCTTC	812
QY	182	TCGCTCTGTTCTTTGCTTAATTCATCCAGCAGTAGTCTAGAGTGGCGGAATACGTCG	241
Db	813	TTGCTCTGTTCTTTGCTTAATTCATCCAGCAGTAGTCTAGAGTGGCGGAATACGTCG	872
QY	242	GCCTCTATGTCCTTACCAACGACTGTTCCATAGCAGTATTGTGTAGAGGCCGATGACG	301
Db	873	GCCTCTATGTCCTTACCAACGACTGTTCCATAGCAGTATTGTGTAGAGGCCGATGACG	932
QY	302	TTATTCTGCACACACCCGGCTGCATACCTCTGTGCCAGGCGCAATACATCCACGTC	361
Db	933	TTATTCTGCACACACCCGGCTGTGTGCTTGTTCAGATGGCAATTAATCCACGTC	992
QY	362	GGACCCCACTGACCTACAGTGGCAGTCAAGTAGTACGTCGGAGCAACCAACCGCTTCG	421

Db	993	GGACTCCAGTGACACCTACAGTGGCAGTGCAGTATGTGCGAGCAACACCGCTTCGATAC	1052
Qy	422	GCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTGCTCTGGCTCTACGTGGGTG	481
Db	1053	GCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTGCTCTGGCTCTACGTGGGTG	1112
Qy	482	ACATGTGTGGGGCTGTCTTCCTCGTGGGACAGCCTTCACGTTTCAGACCTCGTGGCATC	541
Db	1113	ATATGTGTGGGGCTGTCTTCCTCGTGGGACAGCCTTCACGTTTCAGACCTCGTGGCATC	1172
Qy	542	AAACGGTCCAGACCTGTAACTGCTCGCTGTATCCAGGCCATCTTTTACGACATCGAATGG	601
Db	1173	AAACGGTCCAGACCTGTAACTGCTCGCTGTATCCAGGCCATCTTTTACGACATCGAATGG	1232
Qy	602	CTTGGGATATGATGAATCGT	625
Db	1233	CTTGGGATATGATGAATCGT	1256
RESULT 14			
LOCUS	A40621	541 bp	DNA linear PAT 05-MAR-1997
DEFINITION	Sequence 21 from Patent WO9425601.		
ACCESSION	A40621		
VERSION	A40621.1	GI:2296656	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 541)		
AUTHORS	Maertens, G. and Stuyver, L.		
TITLE	NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS THERAPEUTIC AND DIAGNOSTIC AGENTS		
JOURNAL	Patent: WO 9425601-A 21 10-NOV-1994;		
COMMENT	INOGENETICS NV (BE) Other publication CA 2139100 941110 Other publication AU 6722294 941121 Other publication CN 1108030 950906 Other publication FI 946066 941223 Other publication NO 944967 941221 Other publication JP 7508423T 950921.		
FEATURES	Location/Qualifiers		
source	1..541		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
	/clone="BR36-9-20"		
CDS	2..541		
	/note="unnamed protein product; protein sequence is in conflict with the conceptual translation"		
	/codon_start=1		
	/protein_id="CAA02497.1"		
	/db_xref="GI:2296657"		
	/translation="VGAPVGGVARALAHGVRALEDGINPATGNLPGCCSFIFLLALFS CLIPHAASLEWRNTSGLYVLTNDCSNSSIYEADDDVILHTPGCIPCVQDGNSTCWTP VTPVAVKYVGATTASIRSHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQ TVQTCNCSLYPGLSHGRMA"		
BASE COUNT	106 a	156 c	142 g 137 t
ORIGIN			
Query Match 85.6%; Score 539.4; DB 6; Length 541;			
Best Local Similarity 99.8%; Pred. No. 9.7e-140;			
Matches 540; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	63	CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGGCATGCGTGGAGGCCCTTGA	122
Db	1	CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGGCATGCGTGGAGGCCCTTGA	60
Qy	123	AGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCTTCT	182
Db	61	AGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCTTCT	120
Qy	183	CGTCTGTCTCTCTGTTAAATTCATCCAGCAGTAGTACGATGAGTGGCGGGAATACGCTCG	242

Db	121	TGCTCTGTTCTCTTGCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTC	180
Qy	243	CCTCTATGTGCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGCCGATGACGT	302
Db	181	CCTCTATGTGCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGCCGATGACGT	240
Qy	303	TATTTCGACACACCCCGCTGCATACCTTGTGTCCAGGACGGCAATAATCATCCAGTCTG	362
Db	241	TATTTCGACACACCCCGCTGCATACCTTGTGTCCAGGACGGCAATAATCATCCAGTCTG	300
Qy	363	GACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTGCGAGCAACACCGCTTCGATACG	422
Db	301	GACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTGCGAGCAACACCGCTTCGATACG	360
Qy	423	CAGTCATGTGGACCTATTAGTGGCGCGGCCACCAAGATGTGCTCTCGCTCTACGTTGGTGA	482
Db	361	CAGTCATGTGGACCTATTAGTGGCGCGGCCACCAAGATGTGCTCTCGCTCTACGTTGGTGA	420
Qy	483	CATGTGTGGGGCTGTCTTCTCTCGTGGGACAGCCTTCACGTTTCAGACTCTGTCGCCATCA	542
Db	421	CATGTGTGGGGCTGTCTTCTCTCGTGGGACAGCCTTCACGTTTCAGACTCTGTCGCCATCA	480
Qy	543	AACGGTCCAGACCTGTAACTGCTCGCTGTATCCAGGCCATCTTTTTCAGGACATCGAATGGC	602
Db	481	AACGGTCCAGACCTGTAACTGCTCGCTGTATCCAGGCCATCTTTTTCAGGACATCGAATGGC	540
Qy	603	T 603	
Db	541	T 541	
RESULT 15			
LOCUS	AX031599	541 bp	DNA linear PAT 20-SEP-2000
DEFINITION	Sequence 21 from Patent EP1004670.		
ACCESSION	AX031599		
VERSION	AX031599.1	GI:10278836	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1		
AUTHORS	Maertens, G. and Stuyver, L.		
TITLE	Sequences of hepatitis c virus genotypes and their use as therapeutic and diagnostic agents		
JOURNAL	Patent: EP 1004670-A 21 31-MAY-2000;		
	INOGENETICS NV (BE)		
FEATURES	Location/Qualifiers		
source	1..541		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
	/clone="BR36-9-20"		
	<2..>541		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC09675.1"		
	/db_xref="GI:10278837"		
	/translation="VGAPVGGVARALAHGVRALEDGINPATGNLPGCCSFIFLLALFS CLIPHAASLEWRNTSGLYVLTNDCSNSSIYEADDDVILHTPGCIPCVQDGNSTCWTP VTPVAVKYVGATTASIRSHVDLLVGAATMCSALYVGMCGAVFLVGQAFTRPRRHQ TVQTCNCSLYPGLSHGRMA"		
BASE COUNT	106 a	156 c	142 g 137 t
ORIGIN			
Query Match 85.6%; Score 539.4; DB 6; Length 541;			
Best Local Similarity 99.8%; Pred. No. 9.7e-140;			
Matches 540; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	63	CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGGCATGCGTGGAGGCCCTTGA	122
Db	1	CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGGCATGCGTGGAGGCCCTTGA	60
Qy	123	AGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTCTTTTCTATTTTCTTCT	182
Db	61	AGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCTTCT	120
Qy	183	CGTCTGTCTCTCTGTTAAATTCATCCAGCAGTAGTACGATGAGTGGCGGGAATACGCTCG	242


```
QY 123 AGACGGGATAAATTTCGCAACAGGGAATTTGGCCGGTTGCTCCTTTCTATTTTTCCTTCT 182
Db |||||||
QY 61 AGACGGGATAAATTTCGCAACAGGGAATTTGGCCGGTTGCTCCTTTCTATTTTTCCTTCT 120
Db |||||||
QY 183 CGCTCTGTCTCTCTGCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGCG 242
Db |||||||
QY 121 TGCCTGTCTCTCTGCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGCG 180
Db |||||||
QY 243 CCTCTATGTCTCTTACCAACGACTGTTCCAATAGCAGTATTGTACGAGGCCGATGACGT 302
Db |||||||
QY 181 CCTCTATGTCTCTTACCAACGACTGTTCCAATAGCAGTATTGTACGAGGCCGATGACGT 240
Db |||||||
QY 303 TATTCCTGACACACCGGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG 362
Db |||||||
QY 241 TATTCCTGACACACCGGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG 300
Db |||||||
QY 363 GACCCGAGTGACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACCCACCGCTTCGATACG 422
Db |||||||
QY 301 GACCCGAGTGACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACCCACCGCTTCGATACG 360
Db |||||||
QY 423 CAGTCATGTGGACCTATTAGTGGGGCGGCCACGATGTGCTCTGCGCTCTACGTGGGTGA 482
Db |||||||
QY 361 CAGTCATGTGGACCTATTAGTGGGGCGGCCACGATGTGCTCTGCGCTCTACGTGGGTGA 420
Db |||||||
QY 483 CATGTGTGGGGCTGTCTTCTCTGTGGGACAAGCCTTCAGTTTCAGACCTCGTCGCCATCA 542
Db |||||||
QY 421 CATGTGTGGGGCTGTCTTCTCTGTGGGACAAGCCTTCAGTTTCAGACCTCGTCGCCATCA 480
Db |||||||
QY 543 AACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGGC 602
Db |||||||
QY 481 AACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGGC 540
Db |||||||
QY 603 T 603
Db |
QY 541 T 541
Db |
```

Search completed: December 20, 2003, 02:02:07
Job time : 2387.36 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 174.169 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-29
Perfect score: 630
Sequence: 1 ATGGGTAGTCATCGATAC.....TGATGATCACTGGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_13Jun03.*

```
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	100.0	630	17 AAT12965	HCV E1 construct H
2	630	100.0	630	24 AAL48929	Hepatitis C virus
3	536.2	85.1	541	15 AAQ78033	Hepatitis C virus
4	533	84.6	541	15 AAQ78032	Hepatitis C virus
5	504.2	80.0	541	15 AAQ78029	Hepatitis C virus
6	503.2	79.9	540	15 AAQ78031	Hepatitis C virus
7	502.6	79.8	541	15 AAQ78030	Hepatitis C virus
8	499.4	79.3	541	15 AAQ78034	Hepatitis C virus

9	499.4	79.3	541	15	AAQ78035	Hepatitis C virus
10	497.8	79.0	541	15	AAQ78036	Hepatitis C virus
11	383.6	60.9	576	16	AAQ83883	Hepatitis C virus
12	383.6	60.9	576	17	AAT16597	Hepatitis C virus
13	382	60.6	576	16	AAQ83882	Hepatitis C virus
14	382	60.6	576	17	AAT16596	Hepatitis C virus
15	380.4	60.4	576	16	AAQ83881	Hepatitis C virus
16	380.4	60.4	576	17	AAT16595	Hepatitis C virus
17	374	59.4	576	16	AAQ83880	Hepatitis C virus
18	374	59.4	576	17	AAT16594	Hepatitis C virus
19	368.2	58.4	576	16	AAQ83879	Hepatitis C virus
20	368.2	58.4	576	17	AAT16593	Hepatitis C virus
21	344.4	54.7	1734	14	AAQ40435	Hepatitis C virus
22	342.8	54.4	1734	14	AAQ40434	Hepatitis C virus
23	341.2	54.2	1734	14	AAQ40436	Hepatitis C virus
24	339.6	53.9	1734	14	AAQ40425	Hepatitis C virus
25	336	53.3	1734	14	AAQ40437	Hepatitis C virus
26	336	53.3	1734	14	AAQ40438	Hepatitis C virus
27	335.6	53.3	630	17	AAT12966	HCV E1 construct H
28	335.6	53.3	630	24	AAL48930	Hepatitis C virus
29	334.8	53.1	1734	14	AAQ40427	Hepatitis C virus
30	334.8	53.1	9405	14	AAQ40426	Full-length Hepati
31	332.8	52.8	1734	14	AAQ40429	Hepatitis C virus
32	332.8	52.8	3461	15	AAQ64068	Non-A, non-B hepat
33	332.8	52.8	3461	16	AAT30386	5'UTR/CORE/ENV/NS1
34	330	52.4	2116	12	AAQ12242	Encodes PT-NANBH V
35	329.8	52.3	633	17	AAT12706	HCV E1 construct H
36	329.8	52.3	633	24	AAL48915	Hepatitis C virus
37	329.6	52.3	1734	14	AAQ40430	Hepatitis C virus
38	328	52.1	1734	14	AAQ40428	Hepatitis C virus
39	326.8	51.9	932	13	AAQ20923	C10-E12 DNA fragme
40	326.8	51.9	1037	15	AAQ58450	Hepatitis C virus
41	326.8	51.9	1562	19	AAV60672	Fragment #5 isolat
42	326.8	51.9	1953	25	AAL55222	Plasmid pIDK2 DNA
43	326.8	51.9	2033	15	AAQ64913	Hepatitis C virus
44	326.8	51.9	2033	16	AAQ86788	Hepatitis C virus
45	326.8	51.9	2829	19	AAV60673	Fragment #6 isolat

ALIGNMENTS

RESULT 1
AAT12965
ID AAT12965 standard; DNA; 630 BP.
XX
AC AAT12965;
XX
XX 24-SEP-1996 (first entry)
DT
DE HCV E1 construct HCCI62.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
DR WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

181 CTCGCTCTGTTCTCTGCTTAATTCATCCAGCAGCTAGTCTAGTGGCGGAATACGCTT 240
241 GGCCTCTATGTCCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAC 300
241 GGCCTCTATGTCCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAC 300
301 GTTATTCTGCACACACCGGCTGCATACCTGTGTCCAGGACGGCAATACATCCAGTGC 360
301 GTTATTCTGCACACACCGGCTGCATACCTGTGTCCAGGACGGCAATACATCCAGTGC 360
361 TGGACCCAGTCACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACACCGCTTCGATA 420
361 TGGACCCAGTCACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACACCGCTTCGATA 420
421 CGCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTCTCGCTCTACGTGGGT 480
421 CGCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTCTCGCTCTACGTGGGT 480
481 GACATGTGTGGGGCTGTCTTCTCTGTGGGACAAAGCCTTCAGCTTCAGCCCTCGTCCCAT 540
481 GACATGTGTGGGGCTGTCTTCTCTGTGGGACAAAGCCTTCAGCTTCAGCCCTCGTCCCAT 540
541 CAAACGGTCCAGACCTGTAACCTGCTGCTGTACCCAGGCCATCTTTCAGGACATCGAATG 600
541 CAAACGGTCCAGACCTGTAACCTGCTGCTGTACCCAGGCCATCTTTCAGGACATCGAATG 600
601 GCTTGGGATATGATGATGAACCTGGTAATAG 630
601 GCTTGGGATATGATGATGAACCTGGTAATAG 630

RESULT 3

AAQ78033
ID AAQ78033 standard; cDNA; 541 BP.
AC AAQ78033;
XX
DT 25-MAR-2003 (updated)
DT 01-AUG-1995 (first entry)
XX
DE Hepatitis C virus Core/E1 region.
XX
KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.
XX
OS Hepatitis C virus type 3a.
XX
FH Key Location/Qualifiers
FT CDS 2..541
FT /*tag= a
FT /product= Core/E1 polypeptide.
XX
PN WO9425601-A2.
XX
PD 10-NOV-1994.
XX
PF 27-APR-1994; 94WO-EP01323.
XX
PR 27-APR-1993; 93EP-0401099.
PR 05-AUG-1993; 93EP-0402019.
XX
PA (INNO-) INNOGENETICS NV SA.
XX
PI Maertens G, Stuyver L;
XX
DR WPI; 1994-358277/44.
DR P-PSDB; AAR63281.
XX
PT New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates
XX

PS Claim 2; Page 111-112; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous nucleotides selected from an HCV type 3 genomic sequence, more particularly (i) the region spanning positions 417-957 of the Core/E1 region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the region spanning positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to amplify nucleic acid from an isolate belonging to a specific genotype, or as a probe for specific detection/classification of nucleic acid. Polypeptides encoded by the nucleotides in such compositions may be used for immunisation against HCV, for the detection of antibodies directed against HCV and for serotyping. This sequence corresponds to the Core/E1 region of HCV subtype 3a and is taken from a clone designated BR36-9-20.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX

SQ Sequence 541 BP; 106 A; 154 C; 142 G; 139 T; 0 other;

Query Match 85.1%; Score 536.2; DB 15; Length 541;
Best Local Similarity 99.4%; Pred. No. 5.5e-162;
Matches 538; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 63 COTCGCGCTCCCGTAGGAGGCGTCGCAAGAGCCCTTGGCATGGCGTGAGGGCCCTTGA 122
DB 1 COTCGCGCTCCCGTAGGAGGCGTCGCAAGAGCCCTTGGCATGGCGTGAGGGCCCTTGA 60
QY 123 AGACGGGATAAATTTGCGCAACAGGGAATTTGGCCGGTGTCTCTTTTCTATTTCTCTCT 182
DB 61 AGACGGGATAAATTTGCGCAACAGGGAATTTGGCCGGTGTCTCTTTTCTATTTCTCTCT 120
QY 183 CGCTCTGTTCTCTCTGCTTAATTCATCCAGCAGTGTCTAGAGTGCGGGAATACGCTGG 242
DB 121 TGCTCTGTTCTCTCTGCTTAATTCATCCAGCAGTGTCTAGAGTGCGGGAATACGCTGG 180
QY 243 CCTCTATGTCCTTACCAACGACTGTTCCAATAGCAGTATTGTGTACGAGCCGATGACGT 302
DB 181 CCTCTATGTCCTTACCAACGACTGTTCCAATAGCAGTATTGTGTACGAGCCGATGACGT 240
QY 303 TATTTCTGCACACACCGCGCTGCATACCTTGTGTCCAGGACGGCAATACATCCAGTGTG 362
DB 241 TATTTCTGCACACACCGCGCTGTATACCTTGTGTCCAGGACGGCAATACATCCAGTGTG 300
QY 363 GACCCAGTGACACCTACAGTGGGAGTCAAGTACGTCGGAGCAACACCGCTTCGATAGC 422
DB 301 GACCCAGTGACACCTACAGTGGGAGTCAAGTACGTCGGAGCAACACCGCTTCGATAGC 360
QY 423 CAGTCATGTGGACCTATTAGTGGGCGCGCCACGATGTCTCTGGCTCTACGTGGGTGA 482
DB 361 CAGTCATGTGGACCTATTAGTGGGCGCGCCACGATGTCTCTGGCTCTACGTGGGTGA 420
QY 483 CATGTGTGGGCTGTCTTCTCTCGTGGGCAAGCCCTTCAAGTTTCAGACCTCTCGGCCATCA 542
DB 421 CATGTGTGGGCTGTCTTCTCTCGTGGGCAAGCCCTTCAAGTTTCAGACCTCTCGGCCATCA 480
QY 543 AACGGTCCAGACCTGTAACTGCTGTACCCAGGCCATCTTTCAGGACATCGAATGGC 602
DB 481 AACGGTCCAGACCTGTAACTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGGC 540
QY 603 T 603
DB 541 T 541

RESULT 4

AAQ78032
ID AAQ78032 standard; cDNA; 541 BP.
XX
AC AAQ78032;
XX

DT	25-MAR-2003 (updated)	243	CCTCTATGTCCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGACGT	302
DT	01-AUG-1995 (first entry)	181	CTCTATGTCCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGACGT	240
XX				
DE	Hepatitis C virus Core/E1 region.	303	TATTTCTGCACACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG	362
XX				
KW	Hepatitis C virus; HCV; primer; probe; detection; diagnosis;	241	TATTTCTGCACACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG	300
KW	classification; immunisation; prophylaxis; serotyping; ss.			
XX		363	GACCCAGTGACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACCGCTTCGATACG	422
OS	Hepatitis C virus type 3a.	301	GACCCAGTGACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACCGCTTCGATACG	360
XX				
FT	Key	423	CAGTCATGTGCACCTATTAGTGGCGCGGCCACGATGTGCTGTGCGCTTACGTGGGTGA	482
FT	Location/Qualifiers			
FT	2..541			
FT	/*tag= a	361	CAGTCATGTGCACCTATTAGTGGCGCGGCCACGATGTGCTCAGCGCTTCTACGTGGGTGA	420
FT	/product= Core/E1 polypeptide.			
XX				
XX	WO9425601-A2.	483	CATGTGTGGGGCTGTCTTCTCTGCGGACCAAGCTTCAGTTCAGACCTCGTCGCCATCA	542
XX				
XX	10-NOV-1994.	421	TATGTGTGGGGCGCTCTTCTCTGCGGACCAAGCTTCAGTTCAGACCTCGTCGCCATCA	480
XX				
XX		543	AACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGGC	602
XX				
XX		481	AACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGGC	540
XX	(INNO-) INNOGENETICS NV SA.			
XX		603	T 603	
XX				
XX	Maertens G, Stuyver L;	541	T 541	
XX				
XX	WPI; 1994-358277/44.			
XX	P-PSDB; AAR63280.			
XX				
XX	New polynucleotide sequences from hepatitis C virus - and related			
XX	vectors, polypeptide(s) and antibodies, useful for immunisation,			
XX	treatment, diagnosis and typing of HCV isolates			
XX	Claim 2; Page 109-110; 404pp; English.			
XX				
XX	Compositions comprising at least 5, and pref. 8 or more contiguous			
XX	nucleotides selected from an HCV type 3 genomic sequence, more			
XX	particularly (i) the region spanning positions 417-957 of the			
XX	Core/E1 region of HCV subtype 3a; (ii) the region spanning positions			
XX	4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning			
XX	positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the			
XX	region spanning positions 8023-8235 of the NS5 region of the BR36			
XX	subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic			
XX	sequence, may be used as primers to amplify nucleic acid from an			
XX	isolate belonging to a specific genotype, or as a probe for specific			
XX	detection/classification of nucleic acid. Polypeptides encoded by			
XX	the nucleotides in such compositions may be used for immunisation			
XX	against HCV, for the detection of antibodies directed against HCV			
XX	and for serotyping. This sequence corresponds to the Core/E1			
XX	region of HCV subtype 3a and is taken from a clone designated			
XX	BR36-9-13.			
XX	(Updated on 25-MAR-2003 to correct PN field.)			
XX				
XX	Sequence 541 BP; 107 A; 155 C; 142 G; 137 T; 0 other;			
XX				
XX	Query Match 84.6%; Score 533; DB 15; Length 541;			
XX	Best Local Similarity 99.1%; Pred. No. 5.9e-161;			
XX	Matches 536; Conservative 0; Mismatches 5; Indels 0; Gaps 0;			
QY	63	CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGGCGATGCGGTGAGGGCCCTTGA	122	
DB	1	CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGGCGATGCGGTGAGGGCCCTTGA	60	
QY	123	AGACGGGATAAATTTCCGAACAGGAATTTGCCGGTTGCTCTTTCTATTTTCTTCT	182	
DB	61	AGACGGGATAAATTTCCGAACAGGAATTTGCCGGTTGCTCTTTCTATTTTCTTCT	120	
QY	183	CGCTCTGTCTCTGCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATAGTCTGG	242	
DB	121	TGCTCTGTCTCTGCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATAGTCTGG	180	

New polynucleotide sequences from hepatitis C virus - and related vectors, polypeptide(s) and antibodies, useful for immunisation, treatment, diagnosis and typing of HCV isolates

Claim 2; Page 103-104; 404pp; English.

Compositions comprising at least 5, and pref. 8 or more contiguous

CC nucleotides selected from an HCV type 3 genomic sequence, more particularly (i) the region spanning positions 417-957 of the Core/E1 region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the region spanning positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to amplify nucleic acid from an isolate belonging to a specific genotype, or as a probe for specific detection/classification of nucleic acid. Polypeptides encoded by the nucleotides in such compositions may be used for immunisation against HCV, for the detection of antibodies directed against HCV and for serotyping. This sequence corresponds to the Core/E1 region of HCV subtype 3a and is taken from a clone designated HD10-2-5.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 541 BP; 104 A; 153 C; 145 G; 139 T; 0 other;

Query Match 80.0%; Score 504.2; DB 15; Length 541;

Best Local Similarity 95.7%; Pred. No. 1.1e-151;

Matches 518; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 63 CQTGCGCGTCCGTAGGAGCGTGCAGAGCCCTTGCGCATGCGTGAGGGCCCTTGA 122

DB 1 CQTGCGCGTCCGTAGGAGCGTGCAGAGCCCTTGCGCATGCGTGAGGGCCCTTGA 60

QY 123 AGACGGGATAAATTCGCAACAGGAATTTGCCGGTGTCTCTTTCTATTTCTCTCT 182

DB 61 AGACGGGATAAATTCGCAACAGGAATTTGCCGGTGTCTCTTTCTATTTCTCTCT 120

QY 183 CGCTCTGTCTCTTGTCTTAAATTCATCCAGCAGTGTCTAGAGTGCGGAAATACGTCGG 242

DB 121 TGTCTGTCTCTTGTCTTAAATTCATCCAGCAGTGTCTAGAGTGCGGAAATACGTCGG 180

QY 243 CCTCTATGTCCTTACCAAGCTGTTCATATAGCAGTATTTGTAGAGCGCGATGACGT 302

DB 181 CCTCTATGTCCTTACCAAGCTGTTCATATAGCAGTATTTGTAGAGCGCGATGACGT 240

QY 303 TATTTCTGCACACACCGGCTGCATACCTTGTCTCAGGACGCAATATACATCCACGTGCTG 362

DB 241 TATTTCTGCACACACCGGCTGTACCTTGTCTCAGGACGCAATATACATCTGCGTCTG 300

QY 363 GACCCAGTGACCTACAGTGGCAGTCAAGTACGTCGAGCAACACCGCTTCGATAGC 422

DB 301 GACCCAGTGACCTACAGTGGCAGTCAAGTACGTCGAGCAACACCGCTTCGATAGC 360

QY 423 CAGTCATGTGACCTATTAGTGGCGCGCCACGATGCTCTGCGCTCTACGTGGGTGA 482

DB 361 CAGGCATGTAGACATGTTGGTGGCGCGCCACGATGCTCTCTCTACGTGGGTGA 420

QY 483 CATGTGTGGGCTGTCTTCTCTGTGGCAAGCCTTTCAGCTTCAGACCTCGTCGCCATCA 542

DB 421 TATGTGTGGGCGCTTCTCTCTGTGGCAAGCCTTTCAGCTTCAGACCTCGTCGCCATCA 480

QY 543 AACGTCACAGCTGTAACTGCTGCTGTCACCGAGCCATCTTTACGAGCATCGAATGCG 602

DB 481 AACGTCACAGCTGTAACTGCTGCTGTCACCGAGCCATCTTTACGAGCATCGAATGCG 540

QY 603 T 603

DB 541 T 541

RESULT 6

AAQ78031

ID AAQ78031 standard; cDNA; 540 BP.

XX AAQ78031;

XX AAQ78031;

DT 25-MAR-2003 (updated)

DT 21-JUL-1995 (first entry)

XX

DE Hepatitis C virus Core/E1 region.

XX

KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis; classification; immunisation; prophylaxis; serotyping; ss.

XX

OS Hepatitis C virus type 3a.

XX

FH Key Location/Qualifiers

FT CDS 2..541

FT /*tag= a

FT /product= Core/E1 polypeptide.

XX

PN WO9425601-A2.

XX

PD 10-NOV-1994.

XX

PF 27-APR-1994; 94WO-EP01323.

XX

PR 27-APR-1993; 93BP-0401099.

PR 05-AUG-1993; 93EP-0402019.

XX

XX (INNO-) INNOGENETICS NV SA.

XX

PI Maertens G, Stuyver L;

XX

XX WPI; 1994-358277/44.

DR P-PSDB; AAR63279.

XX

XX New polynucleotide sequences from hepatitis C virus - and related vectors, polypeptide(s) and antibodies, useful for immunisation, treatment, diagnosis and typing of HCV isolates

XX

XX Claim 2; Page 107-108; 404pp; English.

XX

CC Compositions comprising at least 5, and pref. 8 or more contiguous nucleotides selected from an HCV type 3 genomic sequence, more particularly (i) the region spanning positions 417-957 of the Core/E1 region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the region spanning positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to amplify nucleic acid from an isolate belonging to a specific genotype, or as a probe for specific detection/classification of nucleic acid. Polypeptides encoded by the nucleotides in such compositions may be used for immunisation against HCV, for the detection of antibodies directed against HCV and for serotyping. This sequence corresponds to the Core/E1 region of HCV subtype 3a and is taken from a clone designated HD10-2-21.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 540 BP; 105 A; 153 C; 144 G; 138 T; 0 other;

Query Match 79.9%; Score 503.2; DB 15; Length 540;

Best Local Similarity 95.7%; Pred. No. 2.3e-151;

Matches 517; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 64 GTCGCGCTCCGTAGGAGCGTGCAGAGCCCTTGCGCATGCGTGAGGGCCCTTGA 123

DB 1 GTCGCGCTCCGTAGGAGCGTGCAGAGCCCTTGCGCATGCGTGAGGGCCCTTGA 60

QY 124 GACGGGATAAATTCGCAACAGGAATTTGCCGGTGTCTCTTTCTATTTCTCTCTC 183

DB 61 GACGGGATAAATTCGCAACAGGAATTTGCCGGTGTCTCTTTCTATTTCTCTCTCT 120

QY 184 GCTCTGTTCTCTTGTCTTAAATTCATCCAGCAGTGTCTAGAGTGGCGGAATACGTCTGC 243

DB 121 GCTCTGTTCTCTTGTCTTAAATTCATCCAGCAGTGTCTAGAGTGGCGGAATACGTCTGC 180

QY 244 CTCTATGTCCTTACCAACGACTGTGTCCTTCAATAGCAGTATTTGTACAGGCCGATGCGTT 303

DB 181 CTCTAGTCTCTTACCAACGACTGTGTCCTTCAATAGCAGTATTTGTATGAGGCCGATGCGTT 240


```
QY 304 ATTCTGCACACACCGGCTGCATACCTTTGTTCAGGACGGCAATACATCCACGTGCTGG 363
DB 241 ATTCTGCACACACCGGCTGTGTACCTTTGTTCAGGACGGTAATACATCTGCGTGTGG 300
QY 364 ACCCAGTGACCTTACAGTGGCAGTCAAGTACGTTCGAGCAACACCGCTTCGATAGC 423
DB 301 ACCCCAGTGACCTTACAGTGGCAGTCAAGTACGTTCGAGCAACACCGCTTCGATAGC 360
QY 424 AGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTCTCTGCGCTCTACGTGGGTGAC 483
DB 361 AGCAATGTAGACATATTGTGGCGCGGCCAGATGTCTCTGCTCTACGTGGGTGAT 420
QY 484 ATGTGTGGGGCTGTCTTCTCTCTGTGGCAAGACCTTCAGTTTCAGACCTTCGCGCATCA 543
DB 421 ATGTGTGGGGCTGTCTTCTCTCTGTGGCAAGACCTTCAGTTTCAGACCTTCGCGCATCA 480
QY 544 ACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGCT 603
DB 481 ACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGCT 540

RESULT 7
AAQ78030 standard; cDNA; 541 BP.
XX
AC AAQ78030;
XX
XX 25-MAR-2003 (updated)
DT 21-JUL-1995 (first entry)
XX
DE Hepatitis C virus Core/E1 region.
XX
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.
XX
XX Hepatitis C virus type 3a.
XX
XX Key Location/Qualifiers
FH 2..541
FT CDS /*tag= a
FT /product= Core/E1 polypeptide.
XX
XX W09425601-A2.
XX
XX 10-NOV-1994.
XX
XX 27-APR-1994; 94WO-EP01323.
XX
XX 27-APR-1993; 93EP-0401099.
XX 05-AUG-1993; 93EP-0402019.
XX
XX (INNO-) INNOGENETICS NV SA.
XX
XX Maertens G, Stuyver L;
XX
XX WPI; 1994-358277/44.
XX P-PSDB; AAR63278.
XX
XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates
XX
XX Claim 2; Page 105-106; 404pp; English.
XX
XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the
CC Core/E1 region of HCV subtype 3a, (ii) the region spanning positions
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
CC region spanning positions 8023-8235 of the NS5 region of the BR36
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
```

```
CC sequence, may be used as primers to amplify nucleic acid from an
CC isolate belonging to a specific genotype, or as a probe for specific
CC detection/classification of nucleic acid. Polypeptides encoded by
CC the nucleotides in such compositions may be used for immunisation
CC against HCV, for the detection of antibodies directed against HCV
CC and for serotyping. This sequence corresponds to the Core/E1
CC region of HCV subtype 3a and is taken from a clone designated
CC HD10-2-14.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 541 BP; 106 A; 154 C; 143 G; 138 T; 0 other;
Query Match 79.8%; Score 502.6; DB 15; Length 541;
Best Local Similarity 95.6%; Pred. No. 3.5e-151;
Matches 517; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 63 CGTCGGCGCTCCCGTAGGAGCGGTGCGAAGAGCCCTTGGCATGGCGTAGGGCCCTTGA 122
DB 1 CGTCGGCGCTCCCGTAGGAGCGGTGCGAAGAGCCCTTGGCATGGCGTAGGGCCCTTGA 60
QY 123 AGACGGGATAAATTTCCCAACAGGGAATTTGCCCGGTTGCTCTCTTTCTATTTCTCTT 182
DB 61 AGACGGGATAAATTTCCCAACAGGGAATTTGCCCGGTTGCTCTCTTTCTATTTCTTCC 120
QY 183 CGCTCTGTCTTCTTGTCTTAATTCACAGCAGTAGCTAGAGTGGCGGAATACGTCTGG 242
DB 121 TGCTCTGTCTTCTTGTCTTAATTCACAGCAGTAGCTAGAGTGGCGGAACACGTCTGG 180
QY 243 CCTCTATGTCTCTTACCAACGACTGTTCCNATAGCAGTATTGTGTACGAGGCCGATGAGCT 302
DB 181 CCTCTATGTCTCTTACCAACGACTGTTCCNATAGCAGTATTGTGTATGAGGCCGATGAGCT 240
QY 303 TATTCTGCACACACCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG 362
DB 241 TATTCTGCACACACCGGCTGTGTACCTTGTGTTCAGGACGGTAATACATCTGCGTGTG 300
QY 363 GACCCAGTGACACTACAGTGGCAGTCAAGTAGCTCGGAGCAACCCGCTTCGATAGC 422
DB 301 GACCCAGTGACACTACAGTGGCAGTCAAGTAGCTCGGAGCAACCCGCTTCGATAGC 360
QY 423 CAGTCATGTGGACCTATTAGTGGGCGCGCCACAGATGTCTTGGCTCTACGTGGGTGA 482
DB 361 CAGGCATGTAGACATATTGGTGGGCGCGCCACATGTGCTCTGCTCTACGTGGGTGA 420
QY 483 CATGTGTGGGCTGTCTTCTCTCGTGGGCAAGCCCTTCAGCTTCAGACTCTGCTGGCATCA 542
DB 421 TATGTGTGGGCGCTTCTCTCGTGGGCAAGCCCTTCAGCTTCAGACTCTGCTGGCATCA 480
QY 543 AACGGTCCAGACCTGTAACTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGCG 602
DB 481 AACGGTCCAGACCTGTAACTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGCG 540
QY 603 T 603
DB 541 T 541

RESULT 8
AAQ78034
ID AAQ78034 standard; cDNA; 541 BP.
XX
AC AAQ78034;
XX
XX 25-MAR-2003 (updated)
DT 01-AUG-1995 (first entry)
XX
XX Hepatitis C virus Core/E1 region.
DE
XX
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.
XX
XX Hepatitis C virus type 3a.
XX
```


Key Location/Qualifiers
CDS 2..541
/*tag= a
/product= Core/E1 polypeptide.

W09425601-A2.
10-NOV-1994.
27-APR-1994; 94WO-EP01323.
27-APR-1993; 93EP-0401099.
05-AUG-1993; 93EP-0402019.
(INNO-) INNOGENETICS NV SA.
Maertens G, Stuyver L;
WPI: 1994-358277/44.
P-PSDB; AAR63282.

New polynucleotide sequences from hepatitis C virus - and related vectors, polypeptide(s) and antibodies, useful for immunisation, treatment, diagnosis and typing of HCV isolates

Claim 2; Page 113-114; 404pp; English.

Compositions comprising at least 5, and pref. 8 or more contiguous nucleotides selected from an HCV type 3 genomic sequence, more particularly (i) the region spanning positions 417-957 of the Core/E1 region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the region spanning positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to amplify nucleic acid from an isolate belonging to a specific genotype, or as a probe for specific detection/classification of nucleic acid. Polypeptides encoded by the nucleotides in such compositions may be used for immunisation against HCV, for the detection of antibodies directed against HCV and for serotyping. This sequence corresponds to the Core/E1 region of HCV subtype 3a and is taken from a clone designated BR33-1-10.
(Updated on 25-MAR-2003 to correct PN field.)

Sequence 541 BP; 100 A; 157 C; 148 G; 136 T; 0 other;
Query Match 79.3%; Score 499.4; DB 15; Length 541;
Best Local Similarity 95.2%; Pred. No. 3.8e-150;
Matches 515; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

63 CGTCGGCGCTCCGCTAGGAGCGGTGCAAGAGCCCTTGGCGATGGGTGAGGGCCCTTGA 122
|||||
1 CGTCGGCGCTCCGCTAGGAGCGGTGCAAGAGCCCTTGGCGATGGGTGAGGGCCCTTGA 60
|||||

123 AGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTGCTCTTTCTATTCTTCTCT 182
|||||

61 GGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTGCTCTTTCTATTCTTCTCT 120
|||||

183 CGCTCTGTTCTTCTGTTAATTCATCCAGCAGTAGCTAGAGTGCGGGAATACGTCGG 242
|||||

121 TGTCTGTTCTTCTGTTAATTCATCCAGCAGTAGCTAGAGTGCGGGAATACGTCGG 180
|||||

243 CCTCTATGTCCTTACCAAGACTGTTTCCAAATAGCAGTATTGTACGAGCGCATGACCT 302
|||||

181 CCTCTATGTCCTTACCAAGACTGTTTCCAAATAGTATTGTATGAGCGCGATGACCT 240
|||||

303 TATTTCTGCACACACCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG 362
|||||

241 TATTTCTGCACGCGCGGCTGTGTACCTTGTGTCCAGGACGGCAATACATCCACGTG 300
|||||

363 GACCCGAGTGACCTTACAGTGCGAGTCAAGTACGTCGAGGACCAACCGCTTCGATACG 422
|||||

Db 301 GACCCAGTAACACCTACAGTGGCAGTACGTGCGGGCAACACCGCTTCGATACG 360
QY 423 CAGTCATGTGGACCTATTAGTGGCGCGCCACGATGTCTCTGCGCTCTACGTGGGTGA 482
|||||
Db 361 CAGTCATGTGGACCTATTAGTGGCGCGCCACGATGTCTCTGCGCTTACGTGGGTGA 420
|||||

QY 483 CATGTGTGGGCTGTCTTCTCTCGTGGGCAAGCCTTCAGCTTCAGACCTCGTGGCCATCA 542
|||||
Db 421 TATGTGTGGGCGGTCTTCTCTCGTGGGCAAGCCTTCAGCTTCAGACCTCGTGGCCATCA 480
|||||

QY 543 AACGCTCCAGACCTGTAACTGCTCGCTGTACCCAGGCCATCTTTCAGGACATCGAATGGC 602
|||||
Db 481 AACGCTCCAGACCTGTAACTGCTCGCTGTACCCAGGCCATCTTTCAGGACATCGAATGGC 602
|||||

QY 603 T 603
|
Db 541 T 541

RESULT 9
AAQ78035
ID AAQ78035 standard; cDNA; 541 BP.
XX
AC AAQ78035;
XX
DT 25-MAR-2003 (updated)
DT 01-AUG-1995 (first entry)
XX
DE Hepatitis C virus Core/E1 region.
XX
KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.
XX
OS Hepatitis C virus type 3a.
XX
FH Key Location/Qualifiers
FT CDS 2..541
FT /*tag= a
FT /product= Core/E1 polypeptide.

W09425601-A2.
10-NOV-1994.
27-APR-1994; 94WO-EP01323.
27-APR-1993; 93EP-0401099.
05-AUG-1993; 93EP-0402019.
(INNO-) INNOGENETICS NV SA.
Maertens G, Stuyver L;
WPI: 1994-358277/44.
P-PSDB; AAR63283.

New polynucleotide sequences from hepatitis C virus - and related vectors, polypeptide(s) and antibodies, useful for immunisation, treatment, diagnosis and typing of HCV isolates

Claim 2; Page 115-116; 404pp; English.

Compositions comprising at least 5, and pref. 8 or more contiguous nucleotides selected from an HCV type 3 genomic sequence, more particularly (i) the region spanning positions 417-957 of the Core/E1 region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the region spanning positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to amplify nucleic acid from an isolate belonging to a specific genotype, or as a probe for specific detection/classification of nucleic acid. Polypeptides encoded by

/product= Core/E1 polypeptide.

CC the nucleotides in such compositions may be used for immunisation
CC against HCV, for the detection of antibodies directed against HCV
CC and for serotyping. This sequence corresponds to the Core/E1
CC region of HCV subtype 3a and is taken from a clone designated
CC BR33-1-19.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX

SQ Sequence 541 BP; 100 A; 155 C; 148 G; 138 T; 0 other;

Query Match 79.3%; Score 499.4; DB 15; Length 541;
Best Local Similarity 95.2%; Pred. No. 3.8e-150; Indels 0; Gaps 0;
Matches 515; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 63 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGACCCCTTGCATGCGGTGAGGGCCCTTGA 122
DB 1 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGACCCCTTGCATGCGGTGAGGGCCCTTGA 60
QY 123 AGACGGGATAAATTTCCGCAACAGGGAATTTGCCGGTTCCTCTTTCTATTTCTTCT 182
DB 61 GGACGGGATAAATTTCCGCAACAGGGAATTTGCCGGTTCCTCTTTCTATTTCTTCT 120
QY 183 CGCTCTGTTCTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTC 242
DB 121 TGTCTGTTCTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTC 180
QY 243 CCTCTATGTCCTTACCAACGAGCTGTTCCAAATAGCAGTATTTGTACGAGCCGATGACGT 302
DB 181 CCTCTATGTCCTTACCAACGAGCTGTTCCAAATAGCAGTATTTGTATGAGCCGATGACGT 240
QY 303 TATTTCTGCACACACCCCGGCTGCATACCTTTGTTCAGGACCGCAATACATCCACGTGCTG 362
DB 241 TATTTCTGCACACACCCCGGCTGCATACCTTTGTTCAGGACCGCAATACATCCACGTGCTG 300
QY 363 GACCCAGTGCACACCTACAGTGCAGTCAAGTACGTTCGAGGCAACACCCGCTTCGATAGC 422
DB 301 GACCCAGTGCACACCTACAGTGCAGTCAAGTACGTTCGAGGCAACACCCGCTTCGATAGC 360
QY 423 CAGTCATGTGGACCTATTAGTGGCGGCGGACGATGCTGCTGCGCTTTACGTTGGGTGA 482
DB 361 CAGTCATGTGGACCTATTAGTGGCGGCGGACGATGCTGCTGCGCTTTACGTTGGGTGA 420
QY 483 CAGTCATGTGGCGGCTGCTTCTCTGTTGGGACAAAGCCTTCACGTTTCAGACCTCGTCCCATCA 542
DB 421 TATGTGTGGGCGGCTTCTCTGTTGGGACAAAGCCTTCACGTTTCAGACCTCGTCCCATCA 480
QY 543 AACGGTCCAGACCTGTAACTGCTGTGTATACCGAGCCATCTTTACGAGACATCGAATGGC 602
DB 481 AACGGTCCAGACCTGTAACTGCTGTGTATACCGAGCCATCTTTACGAGACATCGAATGGC 540
QY 603 T 603
DB 541 T 541

RESULT 10

AAQ78036
ID AAQ78036 standard; cDNA; 541 BP.

XX AC AAQ78036;

XX 25-MAR-2003 (updated)
DT 01-AUG-1995 (first entry)

XX Hepatitis C virus Core/E1 region.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
XX classification; immunisation; prophylaxis; serotyping; ss.

XX Hepatitis C virus type 3a.

XX Key Location/Qualifiers
FH 2.541
FT CDS /*tag= a

FT
XX W09425601-A2.
XX 10-NOV-1994.
XX 27-APR-1994; 94WO-EP01323.
XX 27-APR-1993; 93EP-0401099.
XX 05-AUG-1993; 93EP-0402019.
XX (INNO-) INNOGENETICS NV SA.
XX Maertens G, Stuyver L;
XX WPI; 1994-358277/44.
XX P-PSDB; AAR63284.
XX New polynucleotide sequences from hepatitis C virus - and related
XX vectors, polypeptide(s) and antibodies, useful for immunisation,
XX treatment, diagnosis and typing of HCV isolates
XX Claim 2; Page 117-118; 404pp; English.
XX Compositions comprising at least 5, and pref. 8 or more contiguous
XX nucleotides selected from an HCV type 3 genomic sequence, more
XX particularly (i) the region spanning positions 417-957 of the
XX Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
XX 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
XX positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
XX region spanning positions 8023-8235 of the NS5 region of the BR36
XX subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
XX sequence, may be used as primers to amplify nucleic acid from an
XX isolate belonging to a specific genotype, or as a probe for specific
XX detection/classification of nucleic acid. Polypeptides encoded by
XX the nucleotides in such compositions may be used for immunisation
XX against HCV, for the detection of antibodies directed against HCV
XX and for serotyping. This sequence corresponds to the Core/E1
XX region of HCV subtype 3a and is taken from a clone designated
XX BR33-1-20.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 541 BP; 100 A; 154 C; 148 G; 139 T; 0 other;

Query Match 79.0%; Score 497.8; DB 15; Length 541;
Best Local Similarity 95.0%; Pred. No. 1.2e-149;
Matches 514; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 63 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGACCCCTTGCATGCGGTGAGGGCCCTTGA 122
DB 1 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGACCCCTTGCATGCGGTGAGGGCCCTTGA 60
QY 123 AGACGGGATAAATTTCCGCAACAGGGAATTTGCCGGTTCCTCTTTCTATTTCTTCT 182
DB 61 GGACGGGATAAATTTCCGCAACAGGGAATTTGCCGGTTCCTCTTTCTATTTCTTCT 120
QY 183 CGCTCTGTTCTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTC 242
DB 121 TGTCTGTTCTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTC 180
QY 243 CCTCTATGTCCTTACCAACGAGCTGTTCCAAATAGCAGTATTTGTACGAGCCGATGACGT 302
DB 181 CCTCTATGTCCTTACCAACGAGCTGTTCCAAATAGCAGTATTTGTATGAGCCGATGACGT 240
QY 303 TATTTCTGCACACACCCCGGCTGCATACCTTTGTTCAGGACCGCAATACATCCACGTGCTG 362
DB 241 TATTTCTGCACACCCCGGCTGCATACCTTTGTTCAGGACCGCAATACATCCACGTGCTG 300
QY 363 GACCCAGTGCACACCTACAGTGCAGTCAAGTACGTTCGAGGCAACACCCGCTTCGATAGC 422
DB 301 GACCCAGTGCACACCTACAGTGCAGTCAAGTACGTTCGAGGCAACACCCGCTTCGATAGC 360
QY 423 CAGTCATGTGGACCTATTAGTGGCGGCGGACGATGCTGCTGCGCTTTACGTTGGGTGA 482


```

Db 361 CAGTCATGTGGACCTGTTAGTAGGCGCGCCAGCATGTCTCTGGCGCTTTACGTGGGTGA 420
QY 483 CATGTGTGGGCTGTCTTCTCTGTGGGCAAGCCTTCAGCTTCAGACCTTCGTGCGCCATCA 542
Db 421 TATGTGTGGGCGCTTCTCTGTGGGCAAGCCTTCAGCTTCAGACCTTCGTGCGCCATCA 480
QY 543 AACGGTCCAGACCTGTAACTGCTGCTGTACCGAGCCATCTTTCAGGACATCGAATGGC 602
Db 481 AACGGTCCAGACCTGTAACTGCTGCTGTACCGAGCCATCTTTCAGGACATCGAATGGC 540
QY 603 T 603
Db 541 T 541

RESULT 11
AAQ83883
ID AAQ83883 standard; cDNA; 576 BP.
XX
AC AAQ83883;
XX
XX 25-MAR-2003 (updated)
DT 19-SEP-1995 (first entry)
XX
DE Hepatitis C virus envelope 1 gene cDNA isolate S54.
XX
XX Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate S54;
KW diagnosis; vaccines; antibodies; antisera; gene inhibition; ss.
XX
OS Hepatitis C virus.
XX
XX Key Location/Qualifiers
FH mat_peptide 1..576
FT /*tag= a
XX
XX WO9501442-A2.
PN
XX 12-JAN-1995.
PD
XX 28-JUN-1994; 94WO-US07320.
PF
XX 29-JUN-1993; 93US-0086428.
PR
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Bukh J, Miller RH, Purcell RH;
XX WPI; 1995-061006/08.
DR P-PSDB; AAR69672.
XX
XX Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
PT and in vaccines
XX
XX Claim 1; Pages 72-73; 186pp; English.
PS
XX AAQ83883 encodes AAR69672 hepatitis C virus (HCV) envelope 1 (E1)
CC protein isolate S54, both can be used for the diagnosis of HCV
CC infection, and in the prodn. of anti-HCV vaccines, antibodies
CC and antisera. The cDNA may also be used to inhibit the expression
CC of the HCV E1 gene.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 576 BP; 110 A; 156 C; 163 G; 147 T; 0 other;
SQ
Query Match 60.9%; Score 383.6; DB 16; Length 576;
Best Local Similarity 96.6%; Pred. No. 6.9e-113;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACGTCTGGCGCTCTATGCTTACCAAGACTGTTCGAATAGCACT 279
Db 1 CTAGAGTGGCGGAATACGTCTGGCGCTCTATGCTTACCAAGACTGTTCGAATAGCACT 60

```

```

QY 280 ATTGTGTACGAGCGCCGATGACGTTATTCTTGCAACACCGCGCTGCATACCTTTGTGCCAG 339
Db 61 ATTGTGTATGAGCGCGATGACGTCATTCTTGCAACACACCGCGCTGTGTACCTTTGTGCCAG 120
QY 340 GACGGCAATACATCCACGCTGCTGGACCCGAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCCACGCTGCTGGACCCGAGTGACACCTACAGTGGCAGTCAAGTACGTC 180
QY 400 GGAGCAACACCGCTTCGATACGACGATCATGTGGACCTATTATTAGTGGCGCGGCCACGATG 459
Db 181 GGAGCAACACCGCTTCGATACGACGATCATGTGGACCTATTATTAGTGGCGCGGCCACGATG 240
QY 460 TGCTCTGGCTCTACGTGGGTGACATGTGTGGGCTGTCTTCTCTGTGGGCAAGCCTTC 519
Db 241 TGCTCTGGCTCTATGTGTGGGTGATATGTGTGGGCGCTCTTCTCTGTGGGCAAGCCTTC 300
QY 520 ACGTTCAGACCTCGTCGCCATCAAAACGGTCCAGACCTGTAACTGCTCGTGTACCCAGGC 579
Db 301 ACGTTCAGACCTCGTCGCCATCAAAACGGTCCAGACCTGTAACTGCTCGTGTACCCAGGC 360
QY 580 CATCTTTCAGGACATCGAATGCTTTGGGATATGATGATGAATGCTGT 625
Db 361 CATCTTTCAGGACATCGAATGCTTTGGGATATGATGATGAATGCTGT 406

RESULT 12
AAT16597
ID AAT16597 standard; cDNA; 576 BP.
XX
AC AAT16597;
XX
XX 30-SEP-1996 (first entry)
DT
DE Hepatitis C virus isolate S54 envelope 1 gene.
XX
KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis; ss.
XX
OS Hepatitis C virus.
XX
XX Key Location/Qualifiers
FH CDS 1..576
FT /*tag= a
FT /product= envelope-1 protein
FT /note= "does not contain start or stop codon"
XX
XX WO9605315-A2.
PN
XX 22-FEB-1996.
PD
XX 15-AUG-1995; 95WO-US10398.
PF
XX 15-AUG-1994; 94US-0290665.
PR
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (USSH ) US SEC DEPT HEALTH.
PA
XX Bukh J, Miller RH, Purcell RH;
XX WPI; 1996-139709/14.
DR P-PSDB; AAR89543.
XX
XX DNA and amino acid sequence of HCV envelope 1 and core proteins -
PT used to determine HCV genotype and as vaccines against HCV infection
XX
XX Claim 1; Page 102; 340pp; English.
PS
XX AAT16559-TJ16609 are cDNAs encoding the E1 (envelope-1) protein of 51 HCV
CC isolates. The isolated sequences are useful for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers
CC are also useful for HCV genotyping. Proteins encoded by the cDNAs
CC can be used in vaccines for immunising against HCV infection. The

```


CC proteins may also be used to detect antibodies against HCV in serum,
 CC saliva, lymphocytes or other mononuclear cells. The antibodies may be
 CC used in the prevention of HCV infection.
 XX
 XX Sequence 576 BP; 110 A; 156 C; 163 G; 147 T; 0 other;
 PS Query Match 60.9%; Score 383.6; DB 17; Length 576;
 CC Best Local Similarity 96.6%; Pred. No. 6.9e-113;
 CC Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 220 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTTACCAACGACTGTTCCAATAGCAGT 279
 Db 1 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTTACCAACGACTGTTCCAATAGCAGT 60
 QY 280 ATTGTGTACGAGCGCGATGACGTTATTCTGCACACACCGCGCTGCATACCTTGTGTCAG 339
 Db 61 ATTGTGTATGAGCGCGATGACGTTATTCTGCACACACCGCGCTGCATACCTTGTGTCAG 120
 QY 340 GACGGCAATACATCCAGTGTGACCGCCAGTGCATACGACCTTACGAGTCAAGTACGTC 399
 Db 121 GACGGCAATACATCCAGTGTGACCGCCAGTGCATACGACCTTACGAGTCAAGTACGTC 180
 QY 400 GGAGCAACCAACCGCTTCGATACGAGTGCATGTCGACCTTATGTCGGCGCGCCACGATG 459
 Db 181 GGAGCAACCAACCGCTTCGATACGAGTGCATGTCGACCTTATGTCGGCGCGCCACGATG 240
 QY 460 TGCTCTGCGCTCTATGTCGGGTGACATGTCGGGCTGTCTTCTCGTGGGCAAGCCCTTC 519
 Db 241 TGCTCTGCGCTCTATGTCGGGTGACATGTCGGGCTGTCTTCTCGTGGGCAAGCCCTTC 300
 QY 520 ACCTTCAGACCTCGTCCGCAATCAACCGGTCCAGACCTGTAACCTGTCGTGACCCAGGC 579
 Db 301 ACCTTCAGACCTCGTCCGCAATCAACCGGTCCAGACCTGTAACCTGTCGTGACCCAGGC 360
 QY 580 CATCTTTTCAGGACATCGAATGCGCTTGGGATATGATGATGAATGGT 625
 Db 361 CATCTTTTCAGGACATCGAATGCGCTTGGGATATGATGATGAATGGT 406
 RESULT 13
 AAQ83882 standard; cdna; 576 BP.
 XX
 XX AAQ83882;
 XX
 XX 25-MAR-2003 (updated)
 DT 19-SEP-1995 (first entry)
 XX
 XX Hepatitis C virus envelope 1 gene cDNA isolate S52.
 DE
 XX Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate S52;
 KW diagnosis; vaccines; antibodies; antisera; gene inhibition; ss.
 XX
 XX Hepatitis C virus.
 OS
 XX Key Location/Qualifiers
 FH mat_peptide 1..576
 FT /*tag= a
 XX
 XX WO9501442-A2.
 PN
 XX 12-JAN-1995.
 PD
 XX 28-JUN-1994; 94WO-US07320.
 PF
 XX 29-JUN-1993; 93US-0086428.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Bukh J, Miller RH, Purcell RH;
 FI
 XX WPI; 1995-061006/08.
 DR P-PSDB; AAR69671.
 DR

XX Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
 PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
 PT and in vaccines
 XX
 XX Claim 1; Page 72; 186pp; English.
 PS
 XX AAQ83882 encodes AAR69671 hepatitis C virus (HCV) envelope 1 (E1)
 CC protein isolate S52, both can be used for the diagnosis of HCV
 CC infection, and in the prodn. of anti-HCV vaccines, antibodies
 CC and antisera. The cDNA may also be used to inhibit the expression
 CC of the HCV E1 gene.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 576 BP; 108 A; 153 C; 166 G; 149 T; 0 other;
 PS Query Match 60.6%; Score 382; DB 16; Length 576;
 CC Best Local Similarity 96.3%; Pred. No. 2.3e-112;
 CC Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 220 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTTACCAACGACTGTTCCAATAGCAGT 279
 Db 1 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTTACCAACGACTGTTCCAATAGCAGT 60
 QY 280 ATTGTGTACGAGCGCGATGACGTTATTCTGCACACACCGCGCTGCATACCTTGTGTCAG 339
 Db 61 ATTGTGTATGAGCGCGATGACGTTATTCTGCACACACCGCGCTGCATACCTTGTGTCAG 120
 QY 340 GACGGCAATACATCCAGTGTGACCGCCAGTGCATACGACCTTACGAGTCAAGTACGTC 399
 Db 121 GACGGCAATACATCCAGTGTGACCGCCAGTGCATACGACCTTACGAGTCAAGTACGTC 180
 QY 400 GGAGCAACCAACCGCTTCGATACGAGTGCATGTCGACCTTATGTCGGCGCGCCACGATG 459
 Db 181 GGAGCAACCAACCGCTTCGATACGAGTGCATGTCGACCTTATGTCGGCGCGCCACGATG 240
 QY 460 TGCTCTGCGCTCTATGTCGGGTGACATGTCGGGCTGTCTTCTCGTGGGCAAGCCCTTC 519
 Db 241 TGCTCTGCGCTCTATGTCGGGTGACATGTCGGGCTGTCTTCTCGTGGGCAAGCCCTTC 300
 QY 520 ACCTTCAGACCTCGTCCGCAATCAACCGGTCCAGACCTGTAACCTGTCGTGACCCAGGC 579
 Db 301 ACCTTCAGACCTCGTCCGCAATCAACCGGTCCAGACCTGTAACCTGTCGTGACCCAGGC 360
 QY 580 CATCTTTTCAGGACATCGAATGCGCTTGGGATATGATGATGAATGGT 625
 Db 361 CATCTTTTCAGGACATCGAATGCGCTTGGGATATGATGATGAATGGT 406
 RESULT 14
 AAT16596
 ID AAT16596 standard; cdna; 576 BP.
 XX
 XX AAT16596;
 XX
 XX 30-SEP-1996 (first entry)
 DT
 XX
 XX Hepatitis C virus isolate S52 envelope 1 gene.
 DE
 XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
 KW hepatitis; ss.
 XX
 XX Hepatitis C virus.
 OS
 XX Key Location/Qualifiers
 FH CDS 1..576
 FT /*tag= a
 FT /product= envelope-1.protein
 FT /note= "does not contain start or stop codon"
 XX
 XX WO9605315-A2.
 XX
 XX 22-FEB-1996.
 PD


```
XX 15-AUG-1995; 95WO-US10398.
PF 15-AUG-1994; 94US-0290665.
PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (USSH ) US SEC DEPT HEALTH.
XX
PI Bukh J, Miller RH, Purcell RH;
XX WPI; 1996-139709/14.
DR P-PSDB; AAR89542.
XX
PT DNA and amino acid sequence of HCV envelope 1 and core proteins -
PT used to determine HCV genotype and as vaccines against HCV infection
XX
PS Claim 1; Page 101-102; 340pp; English.
XX
CC AAT16559-T16609 are cDNAs encoding the E1 (envelope-1) protein of 51 HCV
CC isolates. The isolated sequences are useful for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers
CC are also useful for HCV genotyping. Proteins encoded by the cDNAs
CC can be used in vaccines for immunising against HCV infection. The
CC proteins may also be used to detect antibodies against HCV in serum,
CC saliva, lymphocytes or other mononuclear cells. The antibodies may be
CC used in the prevention of HCV infection.
XX
SQ Sequence 576 BP; 108 A; 153 C; 166 G; 149 T; 0 other;
Query Match 60.6%; Score 382; DB 17; Length 576;
Best Local Similarity 96.3%; Pred. No. 2.3e-112;
Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 220 CTAGAGTGGCGGAATACGCTCGGCCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 279
DB 1 CTAGAGTGGCGGAATACGCTCGGCCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 60
QY 280 ATTGTGTAGAGGCGCGATGACGTTATTTCTGACACACACCCGGCTGCATACCTTGTGTCAG 339
DB 61 ATTGTGTATGAGGCGCGATGACGTTATTTCTGACACACACCCGGCTGCATACCTTGTGTCAG 120
QY 340 GACGCAATACATCCACGCTGTCGACCGCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
DB 121 GACGCAATACATCCACGCTGTCGACCGCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 180
QY 400 GGAGCAACACCGCTTCGATAGCAGTCAATAGTGGACCTATTAGTGGGCGCGCCACGATG 459
DB 181 GGAGCAACACCGCTTCGATAGCAGTCAATAGTGGACCTATTAGTGGGCGCGCCACGATG 240
QY 460 TGCTCTGGGCTCTACGTTGGGTGACATGTGTGGGGTGTCTTCTCTGTGGGACAAGCCTTC 519
DB 241 TGCTCTGGGCTCTACGTTGGGTGACATGTGTGGGGTGTCTTCTCTGTGGGACAAGCCTTC 300
QY 520 ACGTTCAGACCTTCGCCCATCAACGGTCCAGACCTGTAACCTGCTCGTGTACCCAGGC 579
DB 301 ACGTTCAGACCTTCGCCCATCAACGGTCCAGACCTGTAACCTGCTCGTGTACCCAGGC 360
QY 580 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGCT 625
DB 361 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGCT 406
RESULT 15
AAQ83881
ID AAQ83881 standard; cDNA; 576 BP.
XX
AC AAQ83881;
XX
DT 25-MAR-2003 (updated)
DT 19-SEP-1995 (first entry)
XX
DE Hepatitis C virus envelope 1 gene cDNA isolate S2.
XX
```

```
KW Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate S2;
KW diagnosis; vaccines; antibodies; antisera; gene inhibition; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT mat_peptide 1..576
FT /*tag= a
XX
PN WO9501442-A2.
XX
PD 12-JAN-1995.
XX
PF 28-JUN-1994; 94WO-US07320.
XX
PR 29-JUN-1993; 93US-0086428.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Bukh J, Miller RH, Purcell RH;
XX
DR WPI; 1995-061006/08.
DR P-PSDB; AAR69670.
XX
PT Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
PT and in vaccines
XX
PS Claim 1; Page 71; 186pp; English.
XX
CC AAQ83881 encodes AAR69670 hepatitis C virus (HCV) envelope 1 (E1)
CC protein isolate S2, both can be used for the diagnosis of HCV
CC infection, and in the prodn. of anti-HCV vaccines, antibodies
CC and antisera. The cDNA may also be used to inhibit the expression
CC of the HCV E1 gene.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 576 BP; 109 A; 156 C; 164 G; 147 T; 0 other;
Query Match 60.4%; Score 380.4; DB 16; Length 576;
Best Local Similarity 96.1%; Pred. No. 7.4e-112;
Matches 390; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 220 CTAGAGTGGCGGAATACGCTCGGCCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 279
DB 1 CTAGAGTGGCGGAATACGCTCGGCCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 60
QY 280 ATTGTGTAGAGGCGCGATGACGTTATTTCTGACACACACCCGGCTGCATACCTTGTGTCAG 339
DB 61 ATTGTGTATGAGGCGCGATGACGTTATTTCTGACACACACCTGCTGTGTGTTTCA 120
QY 340 GACGCAATACATCCACGCTGTCGACCGCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
DB 121 GACGCAATACATCCACGCTGTCGACCGCCAGTGACACCTACAGTGGCAGTCAAGTATGTC 180
QY 400 GGAGCAACACCGCTTCGATAGCAGTCAATAGTGGACCTATTAGTGGGCGCGCCACGATG 459
DB 181 GGAGCAACACCGCTTCGATAGCAGTCAATAGTGGACCTATTAGTGGGCGCGCCACATG 240
QY 460 TGCTCTGGGCTCTACGTTGGGTGACATGTGTGGGGTGTCTTCTCTGTGGGACAAGCCTTC 519
DB 241 TGCTCTGGGCTCTACGTTGGGTGACATGTGTGGGGTGTCTTCTCTGTGGGACAAGCCTTC 300
QY 520 ACGTTCAGACCTTCGCCCATCAACGGTCCAGACCTGTAACCTGCTCGTGTACCCAGGC 579
DB 301 ACGTTCAGACCTTCGCCCATCAACGGTCCAGACCTGTAACCTGCTCGTGTACCCAGGC 360
QY 580 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGCT 625
DB 361 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGCT 406
```

Search completed: December 19, 2003, 18:51:23

Job time : 176.169 secs

THIS PAGE BLANK (USP)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1620.71 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-29

Perfect score: 630

Sequence: 1 ATGGGTAAAGTCATCGATAC.....TGATGATGAAGTGGTAATAG 630

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hcc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	53.8	8.5	488	9 AV755731	AV755731
C 2	40.2	6.4	492	9 AV758366	AV758366
C 3	40.2	6.4	633	28 AZ991479	AZ991479 2M0275007
C 4	38.4	6.1	521	10 BE756035	BE756035 210056 MA

C 5	37.8	6.0	423	13	BW161485
C 6	37.8	6.0	701	13	BW152423
C 7	37.4	5.9	1252	12	BI517398
C 8	37	5.9	398	10	BE428299
C 9	36.8	5.8	356	13	BU862543
C 10	36.8	5.8	544	14	CA925476
C 11	36.8	5.8	568	14	CA925020
C 12	36.8	5.8	569	14	CA924833
C 13	36.8	5.8	577	14	CA924327
C 14	36.8	5.8	578	13	BU881720
C 15	36.6	5.8	465	13	BW153281
C 16	36.6	5.8	474	13	BW203948
C 17	36.6	5.8	491	13	BW156686
C 18	36.6	5.8	493	13	BW156832
C 19	36.6	5.8	497	13	BW282803
C 20	36.6	5.8	525	13	BW286366
C 21	36.6	5.8	550	13	BW162791
C 22	36.6	5.8	550	13	BW286224
C 23	36.6	5.8	551	13	BW158031
C 24	36.6	5.8	581	13	BW157153
C 25	36.6	5.8	582	13	BW162569
C 26	36.6	5.8	586	13	BW287549
C 27	36.6	5.8	618	13	BW157565
C 28	36.6	5.8	661	9	AV874950
C 29	36.6	5.8	661	13	BW152433
C 30	36.6	5.8	681	13	BW162273
C 31	36.6	5.8	687	13	BW150128
C 32	36.6	5.8	688	13	BW159151
C 33	36.6	5.8	690	13	BW154364
C 34	36.6	5.8	691	13	BW163099
C 35	36.6	5.8	698	9	AV865274
C 36	36.6	5.8	706	13	BW150559
C 37	36.6	5.8	706	13	BW162414
C 38	36.6	5.8	707	13	BW150580
C 39	36.6	5.8	709	13	BW159160
C 40	36.6	5.8	710	13	BW160242
C 41	36.6	5.8	711	13	BW152595
C 42	36.6	5.8	716	13	BW152623
C 43	36.6	5.8	716	13	BW152657
C 44	36.6	5.8	718	13	BW152737
C 45	36.6	5.8	718	13	BW162324

ALIGNMENTS

RESULT 1
AV755731/c 488 bp mRNA linear EST 19-OCT-2000
LOCUS AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 488)
Gu,J., Zhao,W., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
LOCATION/Qualifiers

FEATURES

Mon Dec 22 13:28:47 2003

```

source
1. 488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfIIA; Site_2: sfIIB"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN

Query Match 8.5%; Score 53.8; DB 9; Length 488;
Best Local Similarity 62.4%; Pred. No. 0.0026;
Matches 118; Conservative 0; Mismatches 67; Indels 4; Gaps 2;

QY 438 ATTAGTGGCGGCGGCACGATGTGCTCTGGCGCTTACGTGGGTGACATGTGTGGGCTGT 497
Db 476 ATGGGTGGTGTACACTCGCTCTGCTCAGCTTCTTACGTGTGGACCTCTCGACGGAGT 417

QY 498 CTTCTCTGTGGGACAGCCTTCACGTTTCAGCTTCAGACCTCGTCCATCAAAACGGTCCAGACCTG 557
Db 416 GATGCTTGCAGTTTCAGCTGATCA---TCTGGCTCAGCACCATGATGTTTGTGCATGATG 360

QY 558 TAACTGCTCGCTGATCCAGGCGCATCTTTTCAGGACATCAATG-GCTTGGGATATGATGA 616
Db 359 CAACTGCTCATCTATCTCTCTGGCGCCATCATCTGACACCGTATGAGCATGGGACATGATGA 300

QY 617 TGAACCTGGT 625
Db 299 TGAACCTGGT 291

RESULT 2
AV758366/c
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 492)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
TITLE Homo sapiens cDNA BM clones
JOURNAL Unpublished
COMMENT Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. 492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfIIA; Site_2: sfIIB"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN

Query Match 6.4%; Score 40.2; DB 9; Length 492;
Best Local Similarity 57.5%; Pred. No. 8.1;
Matches 111; Conservative 0; Mismatches 78; Indels 4; Gaps 2;

QY 434 ACCTATTAGTGGCGGCGGCACGATGTGCTCTGGCGCTTACGTGGGTGACATGTGTGGG 493
Db 481 ACATATGTTGTGATACACACATGCTGTGATCAGCTCACTAGCTGTGTGGACCTCTGCGTTG 422

QY 494 CTGCTTCTCTGTGGGACAGCCTTCAGTTTCAGACCTCGTCCCATCAAAACGGTCCAGA 553
Db 421 GGGTATCGCTTGCAGCCCAA---CTGATTATCTCTCAGCAGCAACATTTGGTTTGTGCAAG 365

QY 554 CTGTAACTGCTCGTGTACCCAGGCCATCTTTTCAGGAC-ATCGAATGGTTGGGATATG 612
Db 364 AATGCAACTGCTCATCTTCTATCTCTGGCTGCATCACTGGACTACAGTATGCGATAGCGTATG 305

QY 613 ATGATCAACTGGT 625
Db 304 ATGATCAACTGGT 292

RESULT 3
AZ991479 633 bp DNA linear GSS 27-APR-2001
LOCUS 2M0275007R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
DEFINITION clone UUGC2M0275007 R, genomic survey sequence.
ACCESSION AZ991479
VERSION AZ991479.1 GI:13862706
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 633)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mamoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0275 row: 0 column: 07
Seq primer: CACACAGGAACACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 633.
location/Qualifiers
1. 633
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0275007"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were

```


QY	352	TCACGTGCTGACCCAGTGCACCTACAGTGCAGTCAAGTACGTGCGAGCAACACC	411
DB	311	ACNGANGTCAGAAAGCAGATCAGCTGTTACTACACCAACAGATGACGACACAG	252
QY	412	GCTTCGATACGACGATCATGTGACCTATTAGTGGGCGGCGACGATGCTCTGGGCTC	471
DB	251	ACTGCTATAGTACGCTTGGGCGCCATCTTAATATTCCCGGTTCCGACGTAACCCGAC	192
QY	472	TAGTGGGTGACATGTGGGCTCTCTCTCGTGGGACAAAGCTTC	519
DB	191	TCCGAGCTTTCCAGGACGCGAGTGTTCGCGAATCGACGACTCC	144
RESULT 6			
BW152423/c			
LOCUS			
DEFINITION			
ACCSSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			
DB			
QY			

Mon Dec 22 13:28:47 2003

```

LOCUS          CA924833          569 bp      mRNA      linear      EST 27-DEC-2002
DEFINITION     MTUTL.P10.B12 Aspen leaf cDNA Library Populus tremuloides cDNA,
                mRNA sequence.
ACCESSION      CA924833
VERSION        CA924833.1      GI:27411763
KEYWORDS       EST.
SOURCE         Populus tremuloides (quaking aspen)
ORGANISM       Populus tremuloides
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
                ; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 569)
AUTHORS        Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
                C-J
TITLE          Expressed sequence tags from Aspen
JOURNAL        Unpublished
COMMENT        Contact: Tsai C-J
                Plant Biotech Research Center
                Michigan Technological University, School of Forest Resources &
                Environmental Science
                1400 Townsend Drive, Houghton, MI 49931-1295, USA
                Tel: 906 487 2914
                Fax: 906 487 2915
                Email: chtsai@mtu.edu.
FEATURES       source
                1..569
                /organism="Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:3693"
                /clone_lib="Aspen leaf cDNA Library"
                /note="Organ: leaf"
BASE COUNT     183 a 181 c 113 g 92 t
ORIGIN
Query Match    5.8%; Score 36.8; DB 14; Length 569;
Best Local Similarity 50.6%; Pred. No. 62;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 169 TCTATTTTCCTTCGCTCTGTTCTTCTGTTAATTCATCCAGCAGTAGCTAGAGTGG 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 TGTATTGCATCTAGTTTAAATGGCTTTGTTATTGCTCAAGCAAGTGTAGGACGTG 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 229 CGGAATAGCTGGCCTCTATGCTTACCAACGACTGTTCCATACAGTATTGTGTAC 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 453 CCCAAGACGCTGGCCTCAATGACCAAGAACCTCATCGCATATGCGGTGTAGGTGGC 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 289 GAGCCGATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTCCAGGACGG 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 393 TTTGCTGGAGTTGGTGTCTGCCAACCTTGGCGGTGTTGCTGGCTCGGTGG 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
LOCUS          CA924327/c          577 bp      mRNA      linear      EST 27-DEC-2002
DEFINITION     MTUTL.P4.D04 Aspen leaf cDNA Library Populus tremuloides cDNA,
                mRNA sequence.
ACCESSION      CA924327
VERSION        CA924327.1      GI:27411257
KEYWORDS       EST.
SOURCE         Populus tremuloides (quaking aspen)
ORGANISM       Populus tremuloides
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
                ; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 577)
AUTHORS        Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
                C-J
TITLE          Expressed sequence tags from Aspen
JOURNAL        Unpublished
COMMENT        Contact: Tsai C-J
                Plant Biotech Research Center
                Michigan Technological University, School of Forest Resources &
                Environmental Science
                1400 Townsend Drive, Houghton, MI 49931-1295, USA
                Tel: 906 487 2914
                Fax: 906 487 2915
                Email: chtsai@mtu.edu.
FEATURES       source
                1..568
                /organism="Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:3693"
                /clone_lib="Aspen leaf cDNA Library"
                /note="Organ: leaf"
BASE COUNT     183 a 181 c 113 g 91 t
ORIGIN
Query Match    5.8%; Score 36.8; DB 14; Length 568;
Best Local Similarity 50.6%; Pred. No. 62;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 169 TCTATTTTCCTTCGCTCTGTTCTTCTGTTAATTCATCCAGCAGTAGCTAGAGTGG 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 TGTATTGCATCTAGTTTAAATGGCTTTGTTATTGCTCAAGCAAGTGTAGGACGTG 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 229 CGGAATAGCTGGCCTCTATGCTTACCAACGACTGTTCCATACAGTATTGTGTAC 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 453 CCCAAGACGCTGGCCTCAATGACCAAGAACCTCATCGCATATGCGGTGTAGGTGGC 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 289 GAGCCGATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTCCAGGACGG 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 393 TTTGCTGGAGTTGGTGTCTGCCAACCTTGGCGGTGTTGCTGGCTCGGTGG 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
LOCUS          CA924833/c
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```


1400 Townsend Drive, Houghton, MI 49931-1295, USA

Tel: 906 487 2914

Fax: 906 487 2915

Email: chtsai@mtu.edu.

Location/Qualifiers

1. 577

/organism="Populus tremuloides"

/mol_type="mRNA"

/db_xref="taxon:3693"

/clone_lib="Aspen leaf cDNA Library"

/note="Organ: leaf"

192 a 175 c 112 g 98 t

BASE COUNT
ORIGIN

Query Match 5.8%; Score 36.8; DB 14; Length 577;

Best Local Similarity 50.6%; Pred. No. 62;

Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 169 TCTATTTTCCTTCGCTCTGCTCTCTGCTTAATTCATCCAGCAGCTAGCTAGAGTGG 228

Db 546 TGTATTGCATCTAGTTTAATTCCTTCTGTTATGCTCAAGCAAGTCTAGGACGTG 487

Qy 229 CGGAATACGCTGGCTCTATGCTTACCAACGACTGTTCCAATAGCAGTATTGTGTAC 288

Db 486 CCCAAGACGCTGCCCAATGACCAAGAACCTCATCGCATATGGCGGTGTAGTGGC 427

Qy 289 GAGGCCGATGAGTTATTCACACACACCCCGCTGCATACCTTGTGTCACGACGG 344

Db 426 TTTGCTGGAGTTGGTGGTCTGCCAAACCTTGGCGGTGTGCTGGCTCGGTGG 371

RESULT 14

BU881720

LOCUS

DEFINITION UM66TF02 Populus flower cDNA library Populus balsamifera subsp.

trichocarpa cDNA 5 prime, mRNA sequence.

ACCESION BU881720

VERSION BU881720.1 GI:24073244

KEYWORDS EST.

SOURCE Populus balsamifera subsp. trichocarpa

ORGANISM Populus balsamifera subsp. trichocarpa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; rosids

; euroside 1; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 578)

Unneberg, P., Bhale Rao, R.R., Jansson, S. and Sterky, F.

The poplar tree transcriptome: Analysis of expressed sequence tags

from multiple libraries

Unpublished

CONTACT: BHALERAO RUPALI R.

Umea Plant Science Center

Department of Plant Physiology

University of Umea, 901 87 Umea, Sweden

Tel: +46 90 786 5279

Fax: +46 90 786 6676

Email: rupali.bhale Rao@plantphys.umu.se.

Location/Qualifiers

1. 578

/organism="Populus balsamifera subsp. trichocarpa"

/mol_type="mRNA"

/sub_species="trichocarpa"

/db_xref="taxon:3694"

/clone_lib="Populus flower cDNA library"

/note="Organ: flower"

121 a 107 c 153 g 197 t

BASE COUNT
ORIGIN

Query Match 5.8%; Score 36.8; DB 13; Length 578;

Best Local Similarity 50.6%; Pred. No. 62;

Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 169 TCTATTTTCCTTCGCTCTGCTCTCTGCTTAATTCATCCAGCAGCTAGCTAGAGTGG 228

Db 546 TGTATTGCATCTAGTTTAATTCCTTCTGTTATGCTCAAGCAAGTCTAGGACGTG 487

Qy 229 CGGAATACGCTGGCTCTATGCTTACCAACGACTGTTCCAATAGCAGTATTGTGTAC 288

Db 486 CCCAAGACGCTGCCCAATGACCAAGAACCTCATCGCATATGGCGGTGTAGTGGC 427

Qy 289 GAGGCCGATGAGTTATTCACACACACCCCGCTGCATACCTTGTGTCACGACGG 344

Db 426 TTTGCTGGAGTTGGTGGTCTGCCAAACCTTGGCGGTGTGCTGGCTCGGTGG 371

Db 81 TGTATTGCATCTAGTTTAAATTTGCTTCTTGTATTGCTCAAGCAAGTCTAGGACGTG 140

Qy 229 CGGAATACGCTGGCTCTATGCTTACCAACGACTGTTCCAATAGCAGTATTGTGTAC 288

Db 141 CCCAAGACGCTGGCTCAATGACCAAAAGAACTTCATCGCATATGGCGGTGTAGTGGC 200

Qy 289 GAGGCCGATGAGTTATTCGACACACACCCCGCTGCATACCTTGTGTCACGACGG 344

Db 201 TTTGCTGGAGTTGGTGGTCTGCCAAACCTTGGCGGTGTGCTGGCTCGGTGG 256

RESULT 15

BU153281/c

LOCUS

DEFINITION

BU153281 Nori Satoh unpublished cDNA library, gonad ciona

intestinalis cDNA clone rcigd020b10 3', mRNA sequence.

ACCESION BU153281

VERSION BU153281.1 GI:24510506

KEYWORDS EST.

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;

Phlebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 465)

Satoh, F., Shin-i, T., Kohara, Y. and Satoh, N.

Expressed genes in Ciona intestinalis (2002c)

Unpublished

CONTACT: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

Location/Qualifiers

1. 465

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="rcigd020b10"

/tissue type="gonad"

/clone_lib="Nori Satoh unpublished cDNA library, gonad"

121 a 109 c 123 g 111 t 1 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 102; Conservative

5.8%; Score 36.6; DB 13; Length 465;

48.3%; Pred. No. 68;

Mismatches 109; Indels 0; Gaps 0;

Qy 292 GCCGATGACGTTATTCTGCACACACCCCGGTGCATACCTTGTTCACGACGCAATACA 351

Db 373 GCAGACGCTGTTATTCTGCGCTCGCTTCCTCGGTACCCGATACCGATGACGTCATTAAC 314

Qy 352 TCCAGGTGCTGGACCCCGAGTGACACCTACAGTGGGAGTCAAGTACGTCGGAGCAACACC 411

Db 313 ACTGACGCTCAAAAAGCAGCAGTACAGCTGTTACTACACCAACCAACAGTACGACCAACACG 254

Qy 412 GCTTCGATACGACGTCATGTGGACCTATTAGTGGCGCGCCAGCATGTGCTCTGCGCTC 471

Db 253 ACTGCCATAGTGAGCGTTGGGCGCCCATCTTTATATTCGCGGGTTCGCCAGGTACCCCGCAT 194

Qy 472 TACGTGGGTGACATGTGTGGGGCTGTCTTCC 502

Db 193 TCCGAGCTTTTCCCTGGAGCGGAGTGTTC 163

Search completed: December 20, 2003, 06:54:55

Job time : 1624.71 secs

THIS PAGE BLANK (USP 10)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 45.4354 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-29
Perfect score: 630
Sequence: 1 ATGGGTAGGTCATCGATAC.....TGATGATCAACTGGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A-COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B-COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A-COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B-COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS-COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	100.0	630	3	US-08-612-973-29
2	630	100.0	630	3	US-08-927-597-29
3	383.6	60.9	576	1	US-08-086-428B-39
4	383.6	60.9	576	2	US-08-468-570-39
5	383.6	60.9	576	2	US-08-290-665A-39
6	383.6	60.9	576	4	US-08-466-601A-39
7	383.6	60.9	576	5	PCT-US95-10398-39
8	382	60.6	576	1	US-08-086-428B-38
9	382	60.6	576	2	US-08-468-570-38
10	382	60.6	576	2	US-08-290-665A-38
11	382	60.6	576	4	US-08-466-601A-38
12	382	60.6	576	5	PCT-US95-10398-38
13	380.4	60.4	576	1	US-08-086-428B-37
14	380.4	60.4	576	2	US-08-468-570-37
15	380.4	60.4	576	2	US-08-290-665A-37
16	380.4	60.4	576	4	US-08-466-601A-37
17	380.4	60.4	576	5	PCT-US95-10398-37
18	374	59.4	576	1	US-08-086-428B-36
19	374	59.4	576	2	US-08-468-570-36
20	374	59.4	576	2	US-08-290-665A-36
21	374	59.4	576	4	US-08-466-601A-36
22	374	59.4	576	5	PCT-US95-10398-36
23	368.2	58.4	576	1	US-08-086-428B-35
24	368.2	58.4	576	2	US-08-468-570-35
25	368.2	58.4	576	2	US-08-290-665A-35
26	368.2	58.4	576	4	US-08-466-601A-35
27	368.2	58.4	576	5	PCT-US95-10398-35

28 335.6 53.3 630 3 US-08-612-973-31 Sequence 31, Appl
29 335.6 53.3 630 3 US-08-927-597-31 Sequence 31, Appl
30 330 52.4 2116 3 US-08-191-160-21 Sequence 21, Appl
31 329.8 52.3 633 3 US-08-612-973-7 Sequence 7, Appl
32 329.8 52.3 633 3 US-08-927-597-7 Sequence 7, Appl
33 326.8 51.9 932 1 US-08-081-072-15 Sequence 15, Appl
34 326.8 51.9 932 1 US-08-449-093A-15 Sequence 15, Appl
35 326.8 51.9 1037 1 US-08-462-195-3 Sequence 3, Appl
36 326.8 51.9 1037 2 US-08-636-883-3 Sequence 3, Appl
37 326.8 51.9 1037 3 US-09-127-829-3 Sequence 3, Appl
38 325.2 51.6 9595 3 US-09-014-416-4 Sequence 4, Appl
39 325.2 51.6 9599 3 US-09-014-416-6 Sequence 6, Appl
40 324.8 51.6 1167 1 US-08-324-977-9 Sequence 9, Appl
41 324.8 51.6 1167 2 US-08-384-616-9 Sequence 9, Appl
42 324.8 51.6 1167 2 US-08-904-686A-9 Sequence 9, Appl
43 324.8 51.6 1167 3 US-09-315-850-9 Sequence 9, Appl
44 324.8 51.6 1499 1 US-08-324-977-3 Sequence 3, Appl
45 324.8 51.6 1499 2 US-08-384-616-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-29
; Sequence 29, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..627
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..624

US-08-612-973-29

Query Match 100.0%; Score 630; DB 3; Length 630;
 Best Local Similarity 100.0%; Pred. No. 1.4e-200;
 Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGGTAGGTACATGATACCTTACGTGCGGATTCGCGGATCTCATGGGTACATCCCG 60
 1 ATGGGTAGGTACATGATACCTTACGTGCGGATTCGCGGATCTCATGGGTACATCCCG 60
 61 CTCGTGCGGCTCCCGTAGGAGCGTCCAGAGAGCCCTTGGCATGCGGTGAGGGCCCTT 120
 61 CTCGTGCGGCTCCCGTAGGAGCGTCCAGAGAGCCCTTGGCATGCGGTGAGGGCCCTT 120
 121 GAAGACGGATAAATTCGCAACAGGGAATTCGCCGGTTCCTCTTTCTATTTCCTT 180
 121 GAAGACGGATAAATTCGCAACAGGGAATTCGCCGGTTCCTCTTTCTATTTCCTT 180
 181 CTCGCTCTGCTCTTCTTGTCTTAATTCATCCAGCAGTCTAGAGTGGCGGAATACGTC 240
 181 CTCGCTCTGCTCTTCTTGTCTTAATTCATCCAGCAGTCTAGAGTGGCGGAATACGTC 240
 241 GGCCTCTATGTCCTTACCAACGACTGTTCCCAATAGCAGTATTTGTACGAGCGCGATGAC 300
 241 GGCCTCTATGTCCTTACCAACGACTGTTCCCAATAGCAGTATTTGTACGAGCGCGATGAC 300
 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGC 360
 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGC 360
 361 TGGACCCAGTGACACTACAGTGGGAGTCCCAATAGCAGTATTTGTACGAGCGCGATGAC 420
 361 TGGACCCAGTGACACTACAGTGGGAGTCCCAATAGCAGTATTTGTACGAGCGCGATGAC 420
 421 CGCAGTCATGTCGACCTTATAGTGGGCGGCGCACGATGTCCTGCGCTCTACGTGGGT 480
 421 CGCAGTCATGTCGACCTTATAGTGGGCGGCGCACGATGTCCTGCGCTCTACGTGGGT 480
 481 GACATGTGGGGCTGTCTTCTCGTGGGACAAAGCCTTTCAGGTCAGACCTCGTCGCCAT 540
 481 GACATGTGGGGCTGTCTTCTCGTGGGACAAAGCCTTTCAGGTCAGACCTCGTCGCCAT 540
 541 CAACGGTCAGACCTCTAACTGCTGCTGTATCCAGGCGCATCTTTTCAGGACATCGAATG 600
 541 CAACGGTCAGACCTCTAACTGCTGCTGTATCCAGGCGCATCTTTTCAGGACATCGAATG 600
 601 GCTTGGGATGATGATGAATCTGTAATAG 630
 601 GCTTGGGATGATGATGAATCTGTAATAG 630

RESULT 2
 US-08-927-597-29
 ; Sequence 29, Application US/08927597
 ; Patent No. 6245503
 ; GENERAL INFORMATION:
 ; APPLICANT: MAERTENS, GEERT
 ; APPLICANT: BOSMAN, FONS
 ; APPLICANT: DE MARTYNOFF, GUY
 ; APPLICANT: BUYS, MARIE-ANGE
 ; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
 ; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
 ; NUMBER OF SEQUENCES: 111
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHUYE P. C.
 ; STREET: 1100 NORTH GLEBE ROAD
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/927,597
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/612,973
 FILING DATE: 11-MAR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: BYRNE, THOMAS E.
 REGISTRATION NUMBER: 32,205
 REFERENCE/DOCKET NUMBER: 1487-10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 29:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 630 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..627
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 1..624
 US-08-927-597-29

Query Match 100.0%; Score 630; DB 3; Length 630;
 Best Local Similarity 100.0%; Pred. No. 1.4e-200;
 Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGGTAGGTACATGATACCTTACGTGCGGATTCGCGGATCTCATGGGTACATCCCG 60
 1 ATGGGTAGGTACATGATACCTTACGTGCGGATTCGCGGATCTCATGGGTACATCCCG 60
 61 CTCGTGCGGCTCCCGTAGGAGCGTCCAGAGAGCCCTTGGCATGCGGTGAGGGCCCTT 120
 61 CTCGTGCGGCTCCCGTAGGAGCGTCCAGAGAGCCCTTGGCATGCGGTGAGGGCCCTT 120
 121 GAAGACGGATAAATTCGCAACAGGGAATTCGCCGGTTCCTCTTTCTATTTCCTT 180
 121 GAAGACGGATAAATTCGCAACAGGGAATTCGCCGGTTCCTCTTTCTATTTCCTT 180
 181 CTCGCTCTGCTCTTCTTGTCTTAATTCATCCAGCAGTCTAGAGTGGCGGAATACGTC 240
 181 CTCGCTCTGCTCTTCTTGTCTTAATTCATCCAGCAGTCTAGAGTGGCGGAATACGTC 240
 241 GGCCTCTATGTCCTTACCAACGACTGTTCCCAATAGCAGTATTTGTACGAGCGCGATGAC 300
 241 GGCCTCTATGTCCTTACCAACGACTGTTCCCAATAGCAGTATTTGTACGAGCGCGATGAC 300
 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGC 360
 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGC 360
 361 TGGACCCAGTGACACTACAGTGGGAGTCCCAATAGCAGTATTTGTACGAGCGCGATGAC 420
 361 TGGACCCAGTGACACTACAGTGGGAGTCCCAATAGCAGTATTTGTACGAGCGCGATGAC 420
 421 CGCAGTCATGTCGACCTTATAGTGGGCGGCGCACGATGTCCTGCGCTCTACGTGGGT 480
 421 CGCAGTCATGTCGACCTTATAGTGGGCGGCGCACGATGTCCTGCGCTCTACGTGGGT 480
 481 GACATGTGGGGCTGTCTTCTCGTGGGACAAAGCCTTTCAGGTCAGACCTCGTCGCCAT 540
 481 GACATGTGGGGCTGTCTTCTCGTGGGACAAAGCCTTTCAGGTCAGACCTCGTCGCCAT 540
 541 CAACGGTCAGACCTCTAACTGCTGCTGTATCCAGGCGCATCTTTTCAGGACATCGAATG 600
 541 CAACGGTCAGACCTCTAACTGCTGCTGTATCCAGGCGCATCTTTTCAGGACATCGAATG 600
 601 GCTTGGGATGATGATGAATCTGTAATAG 630
 601 GCTTGGGATGATGATGAATCTGTAATAG 630

QY 541 CAAACGGTCCAGACCTGTAACTGCTCGCTGCTACCCAGGCCATCTTTTCAGGACATCGAATG 600
Db 541 CAAACGGTCCAGACCTGTAACTGCTCGCTGCTACCCAGGCCATCTTTTCAGGACATCGAATG 600
QY 601 GCTTGGGATATGATGATGAAGTGGTAATAG 630
Db 601 GCTTGGGATATGATGATGAAGTGGTAATAG 630
RESULT 3
US-08-086-428B-39
; Sequence 39, Application US/08086428B
; Patent No. 5514539
; GENERAL INFORMATION:
; APPLICANT: BURKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/086,428B
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S54
US-08-086-428B-39
Query Match 60.9%; Score 383.6; DB 1; Length 576;
Best Local Similarity 96.6%; Pred. No. 2.9e-118;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 220 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTCTTACCAACGACCTGTTCCAAATAGCAGT 279
Db 1 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTCTTACCAACGACCTGTTCCAAATAGCAGT 60
QY 280 ATTGTGTACAGGCCGATACGCTTATCTGCACACACCGGCTGCATCTGTGTGCCAG 339
Db 61 ATTGTGTATGAGCCGATACGCTCATTCTGCACACACCGGCTGTGTACCTGTGTTTCAG 120
QY 340 GACGGCAATATACATCCACGCTGGAGCCCGACCTACAGTGCAGTCAAGTACGTC 399
Db 121 GACGGCAATATACATCCACGCTGGAGCCCGACCTACAGTGCAGTCAAGTACGTC 180

QY 400 GGAGCAACACCGCTTCGATAGCGAGTCATGTGGACCTATTATAGTGGCGCGGCCAGCATG 459
Db 181 GGAGCAACACCGCTTCGATAGCGAGTCATGTGGACCTATTATAGTGGCGCGGCCAGCATG 240
QY 460 TGCCTCGCGCTCTACGTGGGTGACATGTGTGGGGCTGTCTTCTCTGTGGGACAAAGCTTTC 519
Db 241 TGCCTCGCGCTCTATGTGGGTGATGTGTGGGGCGCTCTTCTCTGTGGGACAAAGCTTTC 300
QY 520 ACGTTACAGACCTCGTGGCCATCAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGC 579
Db 301 ACGTTACAGACCTCGTGGCCATCAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGC 360
QY 580 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 625
Db 361 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 406
RESULT 4
US-08-468-570-39
; Sequence 39, Application US/08468570
; Patent No. 5871962
; GENERAL INFORMATION:
; APPLICANT: BURKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,570
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S54
US-08-468-570-39
Query Match 60.9%; Score 383.6; DB 2; Length 576;
Best Local Similarity 96.6%; Pred. No. 2.9e-118;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;


```
/ REFERENCE/DOCKET NUMBER: 2026-4070US2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ TELEX: 421792
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 576 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ ORIGINAL SOURCE:
/ ORGANISM: homosapiens
/ INDIVIDUAL ISOLATE: S54
/ US-08-466-601A-39

Query Match          60.9%; Score 383.6; DB 4; Length 576;
Best Local Similarity 96.6%; Pred. No. 2.9e-118;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 279
Db 1 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 60

QY 280 ATTGTGTACGAGCGGATGACGCTTATTCGCACACACCCGGCTGCATACCTTGTGTCCAG 339
Db 61 ATTGTGTATGAGCGGATGACGCTTATTCGCACACACCCGGCTGTGTATGTTTCCAG 120

QY 340 GACGGCAATACATCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 180

QY 400 GGAGCAACACCGCTTCGATACGCGATGATGGACCTTATTAGTGGCGCGCCAGCATG 459
Db 181 GGAGCAACACCGCTTCGATACGCGATGATGGACCTTATTAGTGGCGCGCCAGCATG 240

QY 460 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGCTGTCTTCTCGTGGGACAAAGCCTTC 519
Db 241 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGCTGTCTTCTCGTGGGACAAAGCCTTC 300

QY 520 ACGTTTCAGACCTCGTGGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 579
Db 301 ACGTTTCAGACCTCGTGGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 360

QY 580 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGATGATGATGATGATGATGAT 625
Db 361 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGATGATGATGATGATGATGATGAT 406

RESULT 7
PCT-US95-10398-39
/ Sequence 39, Application PC/TUS9510398
/ GENERAL INFORMATION:
/ APPLICANT: BUKH, J., MILLER, R.H. AND
/ APPLICANT: PURCELL, R.H.
/ TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
/ TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
/ TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
/ TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
/ TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
/ NUMBER OF SEQUENCES: 263
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: MORGAN & FINNEGAN
/ STREET: 345 PARK AVENUE
/ CITY: NEW YORK
/ STATE: NEW YORK
/ COUNTRY: USA
/ ZIP: 10154
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1
```

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/10398
/ FILING DATE: 15-AUG-1995
/ CLASSIFICATION:
/ PRIOR APPLICATION NUMBER: 08/086,428
/ FILING DATE: 29 JUNE 1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/290/665
/ FILING DATE: 15 AUGUST 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: RICHARD W. BORK
/ REGISTRATION NUMBER: 36,459
/ REFERENCE/DOCKET NUMBER: 2026-4116
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ TELEX: 421792
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 576 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ ORIGINAL SOURCE:
/ ORGANISM: homosapiens
/ INDIVIDUAL ISOLATE: S54
/ PCT-US95-10398-39

Query Match          60.9%; Score 383.6; DB 5; Length 576;
Best Local Similarity 96.6%; Pred. No. 2.9e-118;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 279
Db 1 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 60

QY 280 ATTGTGTACGAGCGGATGACGCTTATTCGCACACACCCGGCTGCATACCTTGTGTCCAG 339
Db 61 ATTGTGTATGAGCGGATGACGCTTATTCGCACACACCCGGCTGTGTATGTTTCCAG 120

QY 340 GACGGCAATACATCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 180

QY 400 GGAGCAACACCGCTTCGATACGCGATGATGGACCTTATTAGTGGCGCGCCAGCATG 459
Db 181 GGAGCAACACCGCTTCGATACGCGATGATGGACCTTATTAGTGGCGCGCCAGCATG 240

QY 460 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGCTGTCTTCTCGTGGGACAAAGCCTTC 519
Db 241 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGCTGTCTTCTCGTGGGACAAAGCCTTC 300

QY 520 ACGTTTCAGACCTCGTGGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 579
Db 301 ACGTTTCAGACCTCGTGGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 360

QY 580 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGATGATGATGATGATGATGATGAT 625
Db 361 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGATGATGATGATGATGATGATGAT 406

RESULT 8
US-08-086-428B-38
/ Sequence 39, Application US/08086428B
/ Patent No. 5514539
/ GENERAL INFORMATION:
/ APPLICANT: BUKH, J., MILLER, R.H. AND
/ APPLICANT: PURCELL, R.H.
/ TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
/ TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
/ TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
/ TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
```


DIAGNOSTIC METHODS AND VACCINES

Patent No. 5871962
 GENERAL INFORMATION:
 APPLICANT: BURKH, J., MILLER, R.H. AND
 APPLICANT: PURCELL, R.H.
 TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
 TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
 TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
 TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
 TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
 NUMBER OF SEQUENCES: 159
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/468,570
 FILING DATE: 6-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/086,428
 FILING DATE: 29-JUN-1993
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: RICHARD W. BORK
 REGISTRATION NUMBER: 36,459
 REFERENCE/DOCKET NUMBER: 2026-4070US1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 576 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: homosapiens
 INDIVIDUAL ISOLATE: S52
 US-08-086-428B-38

Query Match 60.6%; Score 382; DB 1; Length 576;
 Best Local Similarity 96.3%; Pred. No. 1e-117;
 Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 220 CTAGAGTGGCGGATACGTCCTGCTATGCTCTTACCAACGACTGTTCCCAATAGCAGT 279
 Db 1 CTAGAGTGGCGGATACGTCCTGCTATGCTCTTACCAACGACTGTTCCCAATAGCAGT 60
 QY 280 ATTGTGTACGAGCGGATGACGTTATTCTGACACACCCCGCTGATACCTTGTGTCAG 339
 Db 61 ATTGTGTACGAGCGGATGACGTTATTCTGACACACCCCGCTGATACCTTGTGTCAG 120
 QY 340 GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
 Db 121 GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 180
 QY 400 GGAGCAACACCGCTTCGATACGACCTATGAGACCTATTAGTGGGCGGCGCCAGCAGT 459
 Db 181 GGAGCAACACCGCTTCGATACGACCTATGAGACCTATTAGTGGGCGGCGCCAGCAGT 240
 QY 460 TGCTCTGGGCTCTACGTTGGGTGACATGTTGGGGGTGTTCTTCCTCGTGGGCAAGCCTTC 519
 Db 241 TGCTCTGGGCTCTACGTTGGGTGACATGTTGGGGGTGTTCTTCCTCGTGGGCAAGCCTTC 300
 QY 520 AGCTTCAGACCTCGTGGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 579
 Db 301 AGCTTCAGACCTCGTGGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 360
 QY 580 CATCTTTACGACATCGAATGGCTTGGGATATGATGATGAATGGT 625
 Db 361 CATCTTTACGACATCGAATGGCTTGGGATATGATGATGAATGGT 406

RESULT 9
 US-08-468-570-38
 ; Sequence 38, Application US/08468570

Patent No. 5871962
 GENERAL INFORMATION:
 APPLICANT: BURKH, J., MILLER, R.H. AND
 APPLICANT: PURCELL, R.H.
 TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
 TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
 TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
 TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
 TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
 NUMBER OF SEQUENCES: 159
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/086,428B
 FILING DATE: 29-JUN-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: RICHARD W. BORK
 REGISTRATION NUMBER: 36,459
 REFERENCE/DOCKET NUMBER: 2026-4070
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 576 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: homosapiens
 INDIVIDUAL ISOLATE: S52
 US-08-086-428B-38

Query Match 60.6%; Score 382; DB 1; Length 576;
 Best Local Similarity 96.3%; Pred. No. 1e-117;
 Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 220 CTAGAGTGGCGGATACGTCCTGCTATGCTCTTACCAACGACTGTTCCCAATAGCAGT 279
 Db 1 CTAGAGTGGCGGATACGTCCTGCTATGCTCTTACCAACGACTGTTCCCAATAGCAGT 60
 QY 280 ATTGTGTACGAGCGGATGACGTTATTCTGACACACCCCGCTGATACCTTGTGTCAG 339
 Db 61 ATTGTGTACGAGCGGATGACGTTATTCTGACACACCCCGCTGATACCTTGTGTCAG 120
 QY 340 GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
 Db 121 GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 180
 QY 400 GGAGCAACACCGCTTCGATACGACCTATGAGACCTATTAGTGGGCGGCGCCAGCAGT 459
 Db 181 GGAGCAACACCGCTTCGATACGACCTATGAGACCTATTAGTGGGCGGCGCCAGCAGT 240
 QY 460 TGCTCTGGGCTCTACGTTGGGTGACATGTTGGGGGTGTTCTTCCTCGTGGGCAAGCCTTC 519
 Db 241 TGCTCTGGGCTCTACGTTGGGTGACATGTTGGGGGTGTTCTTCCTCGTGGGCAAGCCTTC 300
 QY 520 AGCTTCAGACCTCGTGGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 579
 Db 301 AGCTTCAGACCTCGTGGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 360
 QY 580 CATCTTTACGACATCGAATGGCTTGGGATATGATGATGAATGGT 625
 Db 361 CATCTTTACGACATCGAATGGCTTGGGATATGATGATGAATGGT 406

RESULT 9
 US-08-468-570-38
 ; Sequence 38, Application US/08468570

QY 520 AGCTTCAGACCTCGTCCCATCAACGGTCCAGACCTGTAACCTGCTCGCTGTACCCAGGC 579
Db 301 AGCTTCAGACCTCGTCCCATCAACGGTCCAGACCTGTAACCTGCTCGCTGTACCCAGGC 360
QY 580 CATCTTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 625
Db 361 CATGTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 406

RESULT 10

US-08-290-665A-38
; Sequence 38, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S52
US-08-290-665A-38

Query Match 60.6%; Score 382; DB 2; Length 576;
Best Local Similarity 96.3%; Pred. No. 1e-117;
Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 220 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTTACCAAGACTGTTCCAAATAGCAGT 279
Db 1 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTTACCAAGACTGTTCCAAATAGCAGT 60
QY 280 ATTGTGTACGAGCCGATACGCTTATCTGCAACACCCGGTGCATACCTGTGTCCAG 339
Db 61 ATTGTGTATGAGCCGATACGCTTATCTGCAACACCCGGTGTGTACCTGTGTGTTCAG 120
QY 340 GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGCAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGCAGTCAAGTACGTC 180

QY 400 GGAGCAACCAACCGCTTCGATACGACAGTCAATGAGACCTATTATTAGTGGCGCGCCACGATG 459
Db 181 GGAGCAACCAACCGCTTCGATACGACAGTCAATGAGACCTATTATTAGTGGCGCGCCACGATG 240
QY 460 TCCTCTGGCTCTAGTGGTGACATGTGTGGGGCTGTCTTCTCGTGGGACAAGCCTTC 519
Db 241 TCCTCTGGCTCTAGTGGTGATGTGTGGGGCGCTCTTCTCGTGGGACAAGCCTTC 300
QY 520 ACCTTCAGACCTCGTCCCATCAACGGTCCAGACCTGTAACCTGCTGTACCCAGGC 579
Db 301 ACCTTCAGACCTCGTCCCATCAACGGTCCAGACCTGTAACCTGCTGTACCCAGGC 360
QY 580 CATCTTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 625
Db 361 CATGTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 406

RESULT 11

US-08-466-601A-38
; Sequence 38, Application US/08466601A
; Patent No. 6572864
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,601A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S52
US-08-466-601A-38

Query Match 60.6%; Score 382; DB 4; Length 576;
Best Local Similarity 96.3%; Pred. No. 1e-117;
Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

RESULT 12
PCT-US95-10398-38
; Sequence 38, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BURK, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELE: 421792
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:

RESULT 13
US-08-086-428B-37
; Sequence 37, Application US/08086428B
; Patent No. 5514539
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/086,428B
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK


```

; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S2
US-08-086-428B-37

Query Match          60.4%; Score 380.4; DB 1; Length 576;
Best Local Similarity 96.1%; Pred. No. 3.4e-117;
Matches 390; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACAGTCTGGCCCTCTATGCTTACCAAGACACTGTTCCCAATAGCAGT 279
DB 1 CTAGAGTGGCGGAATACAGTCTGGCCCTCTATGCTTACCAAGACACTGTTCCCAATAGCAGT 60

QY 280 ATTGTGTACAGGCCGATGACGTTATTCTGCACACACCCGGCTGCATACCTTTGTGTCAG 339
DB 61 ATTGTGTATGAGGCCGATGACGTTATTCTGCACACACCTGGCTGTGTACCTTTGTGTCAG 120

QY 340 GACGGCAATACATCCACGCTGGAGCCCGAGTACAGTGCAGTGCAGTCAAGTACGTC 399
DB 121 GACGGTAATACATCCACGCTGGAGCCCGAGTACAGTGCAGTGCAGTCAAGTATGTC 180

QY 400 GGAGCAACACCGCTTCGATAGCAGTCAATGAGACCTATTAGTGGCGCGGCCAGCATG 459
DB 181 GGAGCAACACCGCTTCGATAGCAGTCAATGAGACCTATTAGTGGCGCGGCCAGCATG 240

QY 460 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGGTGTCTTCTCTGTGGGCAAGCCTTC 519
DB 241 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGGTGTCTTCTCTGTGGGCAAGCCTTC 300

QY 520 AGCTTCAGACCTCGTCCCATCAAGGTCAGACCTGTAACCTGCTGCTGACCCAGGC 579
DB 301 AGCTTCAGACCTCGTCCCATCAAGGTCAGACCTGTAACCTGCTGCTGACCCAGGC 360

QY 580 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGTTGGT 625
DB 361 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGTTGGT 406

```

```

RESULT 14
US-08-468-570-37
; Sequence 37, Application US/08468570
; Patent No. 5871962
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,570
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S2
US-08-468-570-37

Query Match          60.4%; Score 380.4; DB 2; Length 576;
Best Local Similarity 96.1%; Pred. No. 3.4e-117;
Matches 390; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACAGTCTGGCCCTCTATGCTTACCAAGACACTGTTCCCAATAGCAGT 279
DB 1 CTAGAGTGGCGGAATACAGTCTGGCCCTCTATGCTTACCAAGACACTGTTCCCAATAGCAGT 60

QY 280 ATTGTGTACAGGCCGATGACGTTATTCTGCACACACCCGGCTGCATACCTTTGTGTCAG 339
DB 61 ATTGTGTATGAGGCCGATGACGTTATTCTGCACACACCTGGCTGTGTACCTTTGTGTCAG 120

QY 340 GACGGCAATACATCCACGCTGGAGCCCGAGTACAGTGCAGTGCAGTCAAGTACGTC 399
DB 121 GACGGTAATACATCCACGCTGGAGCCCGAGTACAGTGCAGTGCAGTCAAGTATGTC 180

QY 400 GGAGCAACACCGCTTCGATAGCAGTCAATGAGACCTATTAGTGGCGCGGCCAGCATG 459
DB 181 GGAGCAACACCGCTTCGATAGCAGTCAATGAGACCTATTAGTGGCGCGGCCAGCATG 240

QY 460 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGGTGTCTTCTCTGTGGGCAAGCCTTC 519
DB 241 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGGTGTCTTCTCTGTGGGCAAGCCTTC 300

QY 520 AGCTTCAGACCTCGTCCCATCAAGGTCAGACCTGTAACCTGCTGCTGACCCAGGC 579
DB 301 AGCTTCAGACCTCGTCCCATCAAGGTCAGACCTGTAACCTGCTGCTGACCCAGGC 360

QY 580 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGTTGGT 625
DB 361 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGTTGGT 406

```

```

RESULT 15
US-08-290-665A-37
; Sequence 37, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE

```


Mon Dec 22 13:28:46 2003

TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES

;
 ; NUMBER OF SEQUENCES: 263
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN & FINNEGAN
 ; STREET: 345 PARK AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WORDPERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/290,665A
 ; FILING DATE: 15-AUG-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: RICHARD W. BORK
 ; REGISTRATION NUMBER: 36,459
 ; REFERENCE/DOCKET NUMBER: 2026-4116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; TELEX: 421792
 ; INFORMATION FOR SEQ ID NO: 37:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 576 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; ORIGINAL SOURCE:
 ; ORGANISM: homomapiens
 ; INDIVIDUAL ISOLATE: S2
 ;
 ; US-08-290-665A-37

Query Match 60.4%; Score 380.4; DB 2; Length 576;
 Best Local Similarity 96.1%; Pred. No. 3.4e-117;
 Matches 390; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 Qy 220 CTAGAGTGGCGGATAGCTCTGGGCTCTATGTCCTTACCAACGACTGTTCCTCAATAGCAGT 279
 Db 1 CTAGAGTGGCGGATAGCTCTGGGCTCTATGTCCTTACCAACGACTGTTCCTCAATAGCAGT 60
 Qy 280 ATTGTGTACGAGCGCGATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTCTCCAG 339
 Db 61 ATTGTGTATGAGCGCGATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTCTCCAG 120
 Qy 340 GACGGCAATACATCCAGCTGTGGACCCAGTACCTACAGTGGCAGTCAAGTACGTC 399
 Db 121 GACGGTAATACATCCAGCTGTGGACCCAGTACCTACAGTGGCAGTCAAGTACGTC 180
 Qy 400 GGAGCAACACCGCTTCGATACGCGATCATGTGACCTATTAGTGGGCGGCGCCACGATG 459
 Db 181 GGAGCAACACCGCTTCGATACGCGATCATGTGACCTATTAGTGGGCGGCGCCACGATG 240
 Qy 460 TGCTCTGGGCTCTACGCTGGGTGACATGTGTGGGGCTGTCTTCTCGTGGGACAAGCCTTC 519
 Db 241 TGCTCTGGGCTCTACGCTGGGTGATATGTGTGGGGCGGCTTCTTCTCGTGGGACAAGCCTTC 300
 Qy 520 ACGTTTCAGACCTCGTCGCCATCAAAAGGTCAGACCTGTAACTGCTGCTGTACCCAGGC 579
 Db 301 ACGTTTCAGACCTCGTCGCCATCAAAAGGTCAGACCTGTAACTGCTGCTGTACCCAGGC 360
 Qy 580 CATCTTTACGACATCGAATGGCTTGGGATATGATGATGAATGGT 625
 Db 361 CATCTTTACGACATCGCATGGCTTGGGATATGATGATGAATGGT 406

Search completed: December 20, 2003, 07:03:16
 Job time : 46.4354 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2385.36 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-31
Perfect score: 630
Sequence: 1 ATGGGTAAAGTCATCATAC.....TGATGATGAACGTGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_cm.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rtd.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	630	100.0	630	6	A48693	A48693 Sequence 31
2	630	100.0	630	6	AR157341	AR157341 Sequence
3	630	100.0	630	6	AX452780	AX452780 Sequence
4	630	100.0	630	6	AX685032	AX685032 Sequence
5	602	95.6	957	14	HPCBE95A	L29577 Hepatitis C
6	602	95.6	959	6	A40649	A40649 Sequence 49
7	602	95.6	959	6	AX031627	AX031627 Sequence
8	602	95.6	959	6	AX031897	AX031897 Sequence
9	602	95.6	959	6	AX032167	AX032167 Sequence
10	602	95.6	959	6	BD172148	BD172148 New sequ
11	600.8	95.4	959	6	A40653	A40653 Sequence 53
12	600.8	95.4	959	6	AX031631	AX031631 Sequence
13	600.8	95.4	959	6	AX031901	AX031901 Sequence
14	600.8	95.4	959	6	AX032171	AX032171 Sequence
15	600.8	95.4	959	6	BD172150	BD172150 New sequ
16	597.2	94.8	959	6	A40651	A40651 Sequence 51
17	597.2	94.8	959	6	AX031629	AX031629 Sequence
18	597.2	94.8	959	6	AX031899	AX031899 Sequence
19	597.2	94.8	959	6	AX032169	AX032169 Sequence
20	597.2	94.8	959	6	BD172149	BD172149 New sequ
21	580	92.1	580	6	A40645	A40645 Sequence 45
22	580	92.1	580	6	A40647	A40647 Sequence 47
23	580	92.1	580	6	AX031623	AX031623 Sequence
24	580	92.1	580	6	AX031625	AX031625 Sequence
25	580	92.1	580	6	AX031893	AX031893 Sequence
26	580	92.1	580	6	AX031895	AX031895 Sequence
27	580	92.1	580	6	AX032163	AX032163 Sequence
28	580	92.1	580	6	AX032165	AX032165 Sequence
29	580	92.1	580	6	BD172146	BD172146 New sequ
30	580	92.1	580	6	BD172147	BD172147 New sequ
31	573.2	91.0	579	6	A40753	A40753 Sequence 15
32	573.2	91.0	579	6	AX031731	AX031731 Sequence
33	573.2	91.0	579	6	AX032001	AX032001 Sequence
34	573.2	91.0	579	6	AX032271	AX032271 Sequence
35	573.2	91.0	579	6	BD172193	BD172193 New sequ
36	573	91.0	578	14	HPCBE95F	L29582 Hepatitis C
37	539.6	85.7	1848	14	D50466	D50466 Hepatitis C
38	536	85.1	9408	14	AF064490	AF064490 Hepatitis C
39	524.8	83.3	9343	14	HCY1480	Y13184 Hepatitis C
40	510.2	81.0	579	6	A40755	A40755 Sequence 15
41	510.2	81.0	579	6	A40795	A40795 Sequence 19
42	510.2	81.0	579	6	AX031733	AX031733 Sequence
43	510.2	81.0	579	6	AX031773	AX031773 Sequence
44	510.2	81.0	579	6	AX032003	AX032003 Sequence
45	510.2	81.0	579	6	AX032043	AX032043 Sequence

ALIGNMENTS

RESULT 1
A48693
LOCUS A48693
DEFINITION Sequence 31 from Patent WO9604385.
ACCESSION A48693
VERSION A48693.1 GI:2302406
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 630)
AUTHORS Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 31 15-FEB-1996;

Mon Dec 22 13:28:49 2003

```
INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES             Location/Qualifiers
     source            1..630
                        /organism="unidentified"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32644"
     CDS                1..627
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CAA03139.1"
                        /db_xref="GI:2302407"
                        /translation="MGKVIDTLTCGFADLMGYIPLVGGPIGVARALAHGRVLEBDGV
NYAIGMLPGCSFIFILASLICTVPASAVPYRNAGSIYHVTNDPCNSIIVVEADNLI
LHAPGVCPCVMTGNVSRKWQIPTPLSAPSLGANVAPLRRAVDYLAGGAALCSALYVG
DACCALFLVGMFTYPRQHAIVQNCNCSIIYSHVIGHRMNDMMNW"
     mat_peptide        1..624
                        /product="unnamed"
BASE COUNT            126 a 171 c 170 g 163 t
ORIGIN
1 atgggttaaggtcatcgataccctaacgtggcgatttcgcccgatctctatggggtatattcccg 60
1 atgggttaaggtcatcgataccctaacgtggcgatttcgcccgatctctatggggtatattcccg 60
61 ctctgtagcgcccccattggggcgctcgcaaggcctctcgcacacgggtgagggctctt 120
61 ctctgtagcgcccccattggggcgctcgcaaggcctctcgcacacgggtgagggctctt 120
121 gagacggggtaaaactatgcaacagggaatttacccggttgctcttctctatctttatt 180
121 gagacggggtaaaactatgcaacagggaatttacccggttgctcttctctatctttatt 180
181 ctgtgctcttctctggtctgacgggttcgagcttcgagcttcgagcttcgagcttcct 240
181 ctgtgctcttctctggtctgacgggttcgagcttcgagcttcgagcttcgagcttcct 240
241 gggatttatctggttaccgaatgattgcccaaaactcttccatagcttatgagcgagataac 300
241 gggatttatctggttaccgaatgattgcccaaaactcttccatagcttatgagcgagataac 300
301 ctgattcctacacgacacgtggttcggttcggttcggttcggttcggttcggttcggttc 360
301 ctgattcctacacgacacgtggttcggttcggttcggttcggttcggttcggttcggttc 360
361 tgggtccaaaattaccgcttacactgtcagcccgagcctcgagcagctcgagcagctcct 420
361 tgggtccaaaattaccgcttacactgtcagcccgagcctcgagcagctcgagcagctcct 420
421 cgagagcgggttgactacactgtcagcccgagcctcgagcagctcgagcagctcctcttt 480
421 cgagagcgggttgactacactgtcagcccgagcctcgagcagctcgagcagctcctcttt 480
481 gacgcgtgtggggcactattcttggtaggcaaaatgttcacctataggcctcgccagcac 540
481 gacgcgtgtggggcactattcttggtaggcaaaatgttcacctataggcctcgccagcac 540
541 gctacgggtgcagaaactgcgaactgttccatttacagtgggcactgttaccggccacggatg 600
541 gctacgggtgcagaaactgcgaactgttccatttacagtgggcactgttaccggccacggatg 600
601 gcatgggatgatgatgaactggttaataag 630
601 gcatgggatgatgatgaactggttaataag 630

LOCUS                 AR157341                630 bp            DNA            linear            PAT 17-OCT-2001
DEFINITION            Sequence 31 from patent US 6245503.
ACCESSION              AR157341
VERSION                AR157341.1   GI:16218275
KEYWORDS
SOURCE                 Unknown.
ORGANISM              Unclassified.
REFERENCE              1 (bases 1 to 630)
AUTHORS               Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
TITLE                 Purified hepatitis C virus envelope proteins for diagnostic and
                     therapeutic use
JOURNAL               Patent: US 6245503-A 31 12-JUN-2001;
                     Location/Qualifiers
FEATURES
     source            1..630
                        /organism="unknown"
BASE COUNT            126 a 171 c 170 g 163 t
ORIGIN
1 atgggttaaggtcatcgataccctaacgtggcgatttcgcccgatctctatggggtatattcccg 60
1 atgggttaaggtcatcgataccctaacgtggcgatttcgcccgatctctatggggtatattcccg 60
61 ctctgtagcgcccccattggggcgctcgcaaggcctctcgcacacgggtgagggctctt 120
61 ctctgtagcgcccccattggggcgctcgcaaggcctctcgcacacgggtgagggctctt 120
121 gagacggggtaaaactatgcaacagggaatttacccggttgctcttctctatctttatt 180
121 gagacggggtaaaactatgcaacagggaatttacccggttgctcttctctatctttatt 180
181 ctgtgctcttctctggtctgacgggttcgagcttcgagcttcgagcttcgagcttcct 240
181 ctgtgctcttctctggtctgacgggttcgagcttcgagcttcgagcttcgagcttcct 240
241 gggatttatctggttaccgaatgattgcccaaaactcttccatagcttatgagcgagataac 300
241 gggatttatctggttaccgaatgattgcccaaaactcttccatagcttatgagcgagataac 300
301 ctgattcctacacgacacgtggttcggttcggttcggttcggttcggttcggttcggttc 360
301 ctgattcctacacgacacgtggttcggttcggttcggttcggttcggttcggttcggttc 360
361 tgggtccaaaattaccgcttacactgtcagcccgagcctcgagcagctcgagcagctcct 420
361 tgggtccaaaattaccgcttacactgtcagcccgagcctcgagcagctcgagcagctcct 420
421 cgagagcgggttgactacactgtcagcccgagcctcgagcagctcgagcagctcctcttt 480
421 cgagagcgggttgactacactgtcagcccgagcctcgagcagctcgagcagctcctcttt 480
481 gacgcgtgtggggcactattcttggtaggcaaaatgttcacctataggcctcgccagcac 540
481 gacgcgtgtggggcactattcttggtaggcaaaatgttcacctataggcctcgccagcac 540
541 gctacgggtgcagaaactgcgaactgttccatttacagtgggcactgttaccggccacggatg 600
541 gctacgggtgcagaaactgcgaactgttccatttacagtgggcactgttaccggccacggatg 600
601 gcatgggatgatgatgaactggttaataag 630
601 gcatgggatgatgatgaactggttaataag 630

RESULT 3
AX452780
LOCUS                 AX452780                630 bp            DNA            linear            PAT 06-JUL-2002
DEFINITION            Sequence 31 from Patent EP1211315.
ACCESSION              AX452780
```



```
VERSION      AX452780.1  GI:21712465
KEYWORDS     Hepatitis C virus
SOURCE       Hepatitis C virus
ORGANISM     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE        Recombinant vectors for producing hcv envelope proteins
JOURNAL      Patent: EP 1211315-A 31 05-JUN-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
              source
                1..630
                /organism="Hepatitis C virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:11103"
                1..627
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD38077.1"
                /db_xref="GI:21712466"
                /translation="MGKVIDTLTCGFADLMGYIPLVGGPIGVVARALAHGVRLVDGV
                NYATGNLPCGSPSIFILALLSCLTVPASAVPYRNASGIYHVNDPCNSIVVEADNLI
                LHAPGCVPCWMTGNVSRWQVITPTLSAPSLGAVTAPLRRAVDYLAGGAALCSALYVG
                DACGALFLVGQMFYRPRHATVQNCNCSIYSHVTHGRMNDMMNW"
              mat_peptide
                1..624
                /product="unnamed"
              BASE COUNT  126 a 171 c 170 g 163 t
              ORIGIN
                100.0%; Score 630; DB 6; Length 630;
                Best Local Similarity 100.0%; Pred. No. 1.9e-170;
                Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
              QY 1 ATGGGTAAGGTATCGATACCCCTAACGTGCGGATTCGCGATTCGCGATTCATGGGTATATCCCG 60
              Db 1 ATGGGTAAGGTATCGATACCCCTAACGTGCGGATTCGCGATTCGCGATTCATGGGTATATCCCG 60
              QY 61 CTCGTAGGCGGCCCCATTTGGGGCGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCCTT 120
              Db 61 CTCGTAGGCGGCCCCATTTGGGGCGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCCTT 120
              QY 121 GAGGACGGGTAACACTATGATACCAAGGAATTTACCGGTTGCTCTTTCTATCTTTATT 180
              Db 121 GAGGACGGGTAACACTATGATACCAAGGAATTTACCGGTTGCTCTTTCTATCTTTATT 180
              QY 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGCGCTCTGCGAGTTCCTTACCGAATGCTCT 240
              Db 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGCGCTCTGCGAGTTCCTTACCGAATGCTCT 240
              QY 241 GGGATTTATCATGTATACCAATGATGCCCAAACTCTTCCATAGTCTATGAGGCAGATAAC 300
              Db 241 GGGATTTATCATGTATACCAATGATGCCCAAACTCTTCCATAGTCTATGAGGCAGATAAC 300
              QY 301 CTGATCCTACAGCACCTGGTTCGCGCTTGTGTGTCATGACAGGTAATGTAGTAGATGC 360
              Db 301 CTGATCCTACAGCACCTGGTTCGCGCTTGTGTGTCATGACAGGTAATGTAGTAGATGC 360
              QY 361 TGGGTCCAAATTTACCCCTACACTGTGACGCCCGAGCTCGGAGCTACCGGTCCTCTT 420
              Db 361 TGGGTCCAAATTTACCCCTACACTGTGACGCCCGAGCTCGGAGCTACCGGTCCTCTT 420
              QY 421 CGGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCCGTTATACGTAGGA 480
              Db 421 CGGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCCGTTATACGTAGGA 480
              QY 481 GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTTCGCCAGCAC 540
              Db 481 GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTTCGCCAGCAC 540
              QY 541 GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACCGGATG 600
              Db 541 GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACCGGATG 600

RESULT 4
AX685032
LOCUS       AX685032
DEFINITION Sequence 31 from Patent WO0205548.
ACCESSION  AX685032
VERSION     AX685032.1  GI:29371437
KEYWORDS   Hepatitis C virus
SOURCE     Hepatitis C virus
ORGANISM   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE  1
AUTHORS    Maertens,G., Bosman,F. and Buyse,M.A.
TITLE      Purified Hepatitis C Virus envelope proteins for diagnostic and
              therapeutic use
JOURNAL    Patent: WO 0205548-A 31 18-JUL-2002;
              INNOGENETICS N.V. (BE)
FEATURES   Location/Qualifiers
              source
                1..630
                /organism="Hepatitis C virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:11103"
                1..627
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD86530.1"
                /db_xref="GI:29371438"
                /translation="MGKVIDTLTCGFADLMGYIPLVGGPIGVVARALAHGVRLVDGV
                NYATGNLPCGSPSIFILALLSCLTVPASAVPYRNASGIYHVNDPCNSIVVEADNLI
                LHAPGCVPCWMTGNVSRWQVITPTLSAPSLGAVTAPLRRAVDYLAGGAALCSALYVG
                DACGALFLVGQMFYRPRHATVQNCNCSIYSHVTHGRMNDMMNW"
              mat_peptide
                1..624
                /product="unnamed"
              BASE COUNT  126 a 171 c 170 g 163 t
              ORIGIN
                100.0%; Score 630; DB 6; Length 630;
                Best Local Similarity 100.0%; Pred. No. 1.9e-170;
                Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
              QY 1 ATGGGTAAGGTATCGATACCCCTAACGTGCGGATTCGCGATTCGCGATTCATGGGTATATCCCG 60
              Db 1 ATGGGTAAGGTATCGATACCCCTAACGTGCGGATTCGCGATTCGCGATTCATGGGTATATCCCG 60
              QY 61 CTCGTAGGCGGCCCCATTTGGGGCGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCCTT 120
              Db 61 CTCGTAGGCGGCCCCATTTGGGGCGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCCTT 120
              QY 121 GAGGACGGGTAACACTATGATACCAAGGAATTTACCGGTTGCTCTTTCTATCTTTATT 180
              Db 121 GAGGACGGGTAACACTATGATACCAAGGAATTTACCGGTTGCTCTTTCTATCTTTATT 180
              QY 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGCGCTCTGCGAGTTCCTTACCGAATGCTCT 240
              Db 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGCGCTCTGCGAGTTCCTTACCGAATGCTCT 240
              QY 241 GGGATTTATCATGTATACCAATGATGCCCAAACTCTTCCATAGTCTATGAGGCAGATAAC 300
              Db 241 GGGATTTATCATGTATACCAATGATGCCCAAACTCTTCCATAGTCTATGAGGCAGATAAC 300
              QY 301 CTGATCCTACAGCACCTGGTTCGCGCTTGTGTGTCATGACAGGTAATGTAGTAGATGC 360
              Db 301 CTGATCCTACAGCACCTGGTTCGCGCTTGTGTGTCATGACAGGTAATGTAGTAGATGC 360
              QY 361 TGGGTCCAAATTTACCCCTACACTGTGACGCCCGAGCTCGGAGCTACCGGTCCTCTT 420
              Db 361 TGGGTCCAAATTTACCCCTACACTGTGACGCCCGAGCTCGGAGCTACCGGTCCTCTT 420
```


ORIGIN

421	CGGAGAGCCGTTGACTACTAGCGGAGGGCTGCCCTCTGCTCCGGTTATACGTAGA	481
QY		
421	CGGAGAGCCGTTGACTACTAGCGGAGGGCTGCCCTCTGCTCCGGTTATACGTAGA	480
Db		
481	GACCGGTGTGGGGCACTATCTTGGTAGGCCAAATGTTACCTATAGGCCTCGCCACAC	540
QY		
481	GACCGGTGTGGGGCACTATCTTGGTAGGCCAAATGTTACCTATAGGCCTCGCCACAC	540
Db		
541	GCTACCGGTGCAGAACTGCAACTGTTTCCATTACAGTGGCCCATGTTACCGGCCACCGGATG	600
QY		
541	GCTACCGGTGCAGAACTGCAACTGTTTCCATTACAGTGGCCCATGTTACCGGCCACCGGATG	600
Db		
601	GCATGGGATATGATGATGAACCTGTTAATAG	630
QY		
601	GCATGGGATATGATGATGAACCTGTTAATAG	630
Db		

RESULT 5			
HPC8E95A			
LOCUS			
DEFINITION	HPCB8E95A	957 bp ss-RNA	linear VRL 04-MAY-1995
	Hepatitis C virus (BE95)	core protein gene, complete cds;	envelope protein (E1) gene, complete cds.
ACCESSION	L29577		
VERSION	L29577.1	GI:476661	
KEYWORDS	core protein; envelope protein.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		

REFERENCE
1 (bases 1 to 957)
AUTHORS
TITLE
regions of the hepatitis C virus type 5a
1108-1314 (1994)

JOURNAL
Biochem. Biophys. Res. Commun. 202 (3), 1308-1314 (1994)
MEDLINE
9438342
PUBMED
7520237
REFERENCE
2 (bases 1 to 957)
AUTHORS
Stuyver, L., van Arnhem, W., Wyseur, A., Hernandez, F., Delaporte, E.
and Maertens, G.
TITLE
Classification of hepatitis C viruses based on phylogenetic
analysis of the envelope 1 and nonstructural 5B regions and
identification of five additional subtypes
Hepatology 1994; 19: 10134-10138

JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 91 (21), 10134-10138 (1994)
MEDLINE	95023999
PUBMED	7524083
COMMENT	Original source text: Hepatitis C virus (individual_isolate BE95) cDNA to genomic RNA.

FEATURES	Location/Qualifiers
source	1. .957

```

source
1. .957
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="BE95"
/dep_of="taxon:11103"

```

```
CDS
<1..957
/codon_start=1
/product="precursor protein"
/protein_id="AAA65788.1"
/db_xref="GI:793911"
```

```

/db_xref="GI:793911"
/translation="MSTNPKPQKTKNTNRPDVKFGGQIVGVYLLPRRGRPM
GVYRTRKTSBESQRRGRQPIPKARQPTGRSWGQGPWPLYANEGAGWAGVLLSPRG
SRPNWGNDRPRKGRNKLIDTLTCGADLMYGVYIPLVGGPIGVARALAHGVYRVLED
GVNATVALLPGCSFIIFALLSCLTIPASAVPMYASGIIHVTDNCPNSISIVYEDN
LILHAGCPVCPVMTGNYSRWQVITPTLSAPSLGATAPLRAVDYLAGGAALCSALI
WQDAACQSLPTVQWQWMTNVCNCSIVGSHVYTHGRMA"

```

```
mat_peptide
1. .573
VGDACGATTFLVGQMFIIKFRQIIR
/product="core protein"
```

```
mat_peptide
574. .957
/product="core protein"
/gene="E1"
/product="envelope protein"
```

```
gene
/product="envelope protein
574. .957
/cene="E1"
```

BASE COUNT	189 a	283 c	270 g	215 b
BASE COUNT	189 a	283 c	270 g	215 b

Query Match	95.6%;	Score 602;	DB 14;	Length 957;
Best Local Similarity	100.0%;	Pred. No. 2.3e-162;		
Matches 602;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2	TGGGTAAAGTCAATCGATACCCATAACGTGCGGATTGCGCGATCTCATGGGTATATCCCCG	61	
DB	356	TGGGTAAAGTCAATCGATACCCATAACGTGCGGATTGCGCGATCTCATGGGTATATCCCCG	415	
QY	62	TCCTAGGCGGCCCCATTGGGGGGCTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCTCTTG	121	
DB	416	TCCTAGGCGGCCCCATTGGGGGGCTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCTCTTG	475	
QY	122	AGGACGGGGTAAACTATGCAACAGGGAATTTACCCGGTTGCTCTTTCTCTATCTTTATTC	181	
DB	476	AGGACGGGGTAAACTATGCAACAGGGAATTTACCCGGTTGCTCTTTCTCTATCTTTATTC	535	
QY	182	TTGCTCTTCTCTCGTGTCTGACGGTTCCGGCTCTGCAAGTTCCTACCGAAATGCGCTCTG	241	
DB	536	TTGCTCTTCTCTCGTGTCTGACGGTTCCGGCTCTGCAAGTTCCTACCGAAATGCGCTCTG	595	
QY	242	GGATTTATCATGTTAACATGATTGCCCAACTCTTCCATAGTCTATGAGGCAGATAACC	301	
DB	596	GGATTTATCATGTTAACCAATGATTGCCCAACTCTTCCATAGTCTATGAGGCAGATAACC	655	
QY	302	TGATCTTACACGCACTCGTGGTTCGGTGTGCTTGTGTGTCATGACAGGTAATGTGAGTAGATGCT	361	
DB	656	TGATCTTACACGCACTGGTTCGGTGTGCTTGTGTGTCATGACAGGTAATGTGAGTAGATGCT	715	
QY	362	GGGTCCAAATTAACCCCTACACTGTGAGCCCGAGGCTCGGAGCAGTCAACGGCTCCTCTTC	421	
DB	716	GGGTCCAAATTAACCCCTACACTGTGAGCCCGAGGCTCGGAGCAGTCAACGGCTCCTCTTC	775	
QY	422	GGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCCCTCTGCTCCGCGTTATACGTAGAGAG	481	
DB	776	GGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCCCTCTGCTCCGCGTTATACGTAGAGAG	835	
QY	482	ACGCGTGTGGGCACTATTCTTGTTAGGCCCAATGTTCCACCTATAGGCTCGCCACGACAG	541	
DB	836	ACGCGTGTGGGCACTATTCTTGTTAGGCCCAATGTTCCACCTATAGGCTCGCCACGACAG	895	
QY	542	CTACCGGTGCAGAACTGCAACTGTTCCATTTCAGTGGGCATGTTTACCGGCCACCGGATGG	601	
DB	896	CTACCGGTGCAGAACTGCAACTGTTCCATTTCAGTGGGCATGTTTACCGGCCACCGGATGG	955	
QY	602	CA 603		
DB	956	CA 957		

RESULT 6

LOCUS	LINEAR	DNA	959 bp
A40649			
SEQUENCE			49 from Patent WO9425601.
ACCESSION			A40649
VERSION			A40649.1 GI:2296684
KEYWORDS			unidentified
SOURCE			unidentified
ORGANISM			unclassified.
REFERENCE			1 (bases 1 to 959)
AUTHORS			Maertens, G. and Stuyver, L.
TITLE			NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS THERAPEUTIC AND DIAGNOSTIC AGENTS
JOURNAL			PATENT: WO 9425601-A 49 10-NOV-1994; INNOGENETICS NV (BE)
COMMENT			Other publication CA 2139100 941110 Other publication AU 6722294 941121 Other publication CN 1108030 950906 Other publication FI 946066 941223 Other publication NO 944967 941221 Other publication JP 7508423T 950921.

mat_peptide

mat peptide

gene

COUNT

BASE COUNT

FEATURES	Location/Qualifiers		Accession	
	source		VERSION	
			KEYWORDS	
			SOURCE	
CDS	1..>959		AX031631.1	
	/organism="unidentified"		GI:10278868	
	/mol_type="genomic DNA"			
	/db_xref="taxon:32644"			
BASE COUNT	2..>959			
	/note="unnamed protein product"			
	/codon_start=2			
	/protein_id="CAA02513.1"			
ORIGIN	188 a 285 c 269 g 214 t 3 others			
	95.4%; Score 600.8; DB 6; Length 959;			
	Best Local Similarity 99.5%; Pred. No. 5.1e-162;			
	Matches 599; Conservative 3; Mismatches 0; Indels 0; Gaps 0;			
Qy	2	TGGGTAAGGTCATCGATACCTAACGTCGGGATTCGCGGATCTCATGGGGTATATCCCGC	61	
Db	358	TGGGTAAGGTCATCGATACCTAACGTCGGGATTCGCGGATCTCATGGGGTATATCCCGC	417	
Qy	62	TCGTAGCGGGCCCATTTGGGGGCGTCCGAAGGCTCTCGCACACGCTGTGAGGGTCTTTG	121	
Db	418	TCGTAGCGGGCCCATTTGGGGGCGTCCGAAGGCTCTCGCACACGCTGTGAGGGTCTTTG	477	
Qy	122	AGGACGGGTAAACTATGCAACAGGGAATTTACCGGTTCCTCTTCTCTATCTTTATTC	181	
Db	478	AGGACGGGTAAACTATGCAACAGGGAATTTACCGGTTCCTCTTCTCTATCTTTATTC	537	
Qy	182	TTGCTCTTCTCTGCTGTGACGCTTCGGGCTTCGAGTTCCCTACCGAAATGCCCTCTG	241	
Db	538	TTGCTCTTCTCTGCTGTGACGCTTCGGGCTTCGAGTTCCCTACCGAAATGCCCTCTG	597	
Qy	242	GGATTATCATGTTACCAATGATGCCAACTCTTCCATAGTCTATGAGCGAGATAACC	301	
Db	598	GGATTATCATGTTACCAATGATGCCAACTCTTCCATAGTCTATGAGCGAGATAACC	657	
Qy	302	TGATCTTACACGACCTGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGG	361	
Db	658	TGATCTTACACGACCTGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGG	717	
Qy	362	GGGTCCAAATTTACCCCTACACTGTGACGCTTCGGGCTTCGGGCTTCGGGCTTCGGG	421	
Db	718	GGGTCCAAATTTACCCCTACACTGTGACGCTTCGGGCTTCGGGCTTCGGGCTTCGGG	777	
Qy	422	GGAGAGCGGTGACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	481	
Db	778	GGAGAGCGGTGACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	837	
Qy	482	ACGCGTGGGGCACTATTCTTGGTAGCCAAATGTTCACTATAGGCTTCGCGAGCAG	541	
Db	838	ACGCGTGGGGCACTATTCTTGGTAGCCAAATGTTCACTATAGGCTTCGCGAGCAG	897	
Qy	542	CTACGCTGACAACTGTCATTTACAGTGGCCATGTTACCGGCCACCGGATGG	601	
Db	898	CTACGCTGACAACTGTCATTTACAGTGGCCATGTTACCGGCCACCGGATGG	957	
Qy	602	CA 603		
Db	958	CA 959		
RESULT 12				
AX031631				
LOCUS				
DEFINITION				
Sequence 53 from Patent EP1004670.				
AX031631 959 bp DNA linear PAT 20-SEP-2000				

FEATURES	Location/Qualifiers		Accession	
	source		VERSION	
			KEYWORDS	
			SOURCE	
CDS	1..>959		AX031631.1	
	/organism="unidentified"		GI:10278868	
	/mol_type="genomic DNA"			
	/db_xref="taxon:32644"			
BASE COUNT	2..>959			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAC09691.1"			
ORIGIN	188 a 285 c 269 g 214 t 3 others			
	95.4%; Score 600.8; DB 6; Length 959;			
	Best Local Similarity 99.5%; Pred. No. 5.1e-162;			
	Matches 599; Conservative 3; Mismatches 0; Indels 0; Gaps 0;			
Qy	2	TGGGTAAGGTCATCGATACCTAACGTCGGGATTCGCGGATCTCATGGGGTATATCCCGC	61	
Db	358	TGGGTAAGGTCATCGATACCTAACGTCGGGATTCGCGGATCTCATGGGGTATATCCCGC	417	
Qy	62	TCGTAGCGGGCCCATTTGGGGGCGTCCGAAGGCTCTCGCACACGCTGTGAGGGTCTTTG	121	
Db	418	TCGTAGCGGGCCCATTTGGGGGCGTCCGAAGGCTCTCGCACACGCTGTGAGGGTCTTTG	477	
Qy	122	AGGACGGGTAAACTATGCAACAGGGAATTTACCGGTTCCTCTTCTCTATCTTTATTC	181	
Db	478	AGGACGGGTAAACTATGCAACAGGGAATTTACCGGTTCCTCTTCTCTATCTTTATTC	537	
Qy	182	TTGCTCTTCTCTGCTGTGACGCTTCGGGCTTCGAGTTCCTTACCGAAATGCCCTCTG	241	
Db	538	TTGCTCTTCTCTGCTGTGACGCTTCGGGCTTCGAGTTCCTTACCGAAATGCCCTCTG	597	
Qy	242	GGATTATCATGTTACCAATGATGCCAACTCTTCCATAGTCTATGAGCGAGATAACC	301	
Db	598	GGATTATCATGTTACCAATGATGCCAACTCTTCCATAGTCTATGAGCGAGATAACC	657	
Qy	302	TGATCTTACACGACCTGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGG	361	
Db	658	TGATCTTACACGACCTGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGG	717	
Qy	362	GGGTCCAAATTTACCCCTACACTGTGACGCTTCGGGCTTCGGGCTTCGGGCTTCGGG	421	
Db	718	GGGTCCAAATTTACCCCTACACTGTGACGCTTCGGGCTTCGGGCTTCGGGCTTCGGG	777	
Qy	422	GGAGAGCGGTGACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	481	
Db	778	GGAGAGCGGTGACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	837	
Qy	482	ACGAGCGGTGACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	897	
Db	838	ACGAGCGGTGACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	897	
Qy	482	ACGAGCGGTGACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	541	
Db	898	ACGAGCGGTGACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	897	
Qy	542	CTACGCTGACAACTGTCATTTACAGTGGCCATGTTACCGGCCACCGGATGG	601	
Db	902	CTACGCTGACAACTGTCATTTACAGTGGCCATGTTACCGGCCACCGGATGG	601	


```

Db      898 CTACGGTGCAGAACTGCAACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATGG 757
QY      602 CA 603
Db      958 CA 959

RESULT 13
LOCUS   AX031901
DEFINITION
Sequence 53 from Patent EP0984068.
ACCESSION AX031901
VERSION   AX031901.1 GI:10279051
KEYWORDS
SOURCE   unidentified
ORGANISM unclassified.
REFERENCE
1. Maertens, G. and Stuyver, L.
Sequences of hepatitis c virus genotypes and their use as
therapeutic and diagnostic agents
Patent: EP 0984068-A 53 08-MAR-2000;
INNOGENETICS NV (BE)
FEATURES
source
Location/Qualifiers
1..959
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/clone="PC C/E1"
3..959
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:10279052"
/translation="MSTNPKPORKYNTNRRPQDKVFPGGQIVGGVLLPRRPRGM
GVRATRKTSERSQPRRRQPIPKARQPTGRSGWQPGYPMPLVYANGLGWAGLLSPRG
SRPNMGPNDPRKSRNLGKVIDLTLCGFADLMGYIPLVGGPXXGVARALAHGVRLSD
GVNYATGNLPGCSFSIFILALLSLCTVPASAVPYRNASGIYHVTNDPCNSSIYVEADN
LILHAPGCVPCVMTGNVSRWCVOITPTLSAPSLGAVTAPLRRAVDYLAGGAALCSALY
VGDACGALPLVQGMFTYRPRQATVQNCNSIYSGHVTGHRMA"
BASE COUNT 188 a 285 c 269 g 214 t 3 others
ORIGIN
Query Match 95.4%; Score 600.8; DB 6; Length 959;
Best Local Similarity 99.5%; Pred. No. 5.1e-162;
Matches 599; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 TGGGTAAGGTCTATCGATACCTAAACGTGCGGATTCGCCGATCTCATGGGGTATATCCCGC 61
Db      358 TGGGTAAGGTCTATCGATACCTAAACGTGCGGATTCGCCGATCTCATGGGGTATATCCCGC 417
QY      62 TCGTAGCGCGGCCCATTTGGGGCGTCCGAAGGGCTCTCGCACACGGGTGTAGGGTCTTGG 121
Db      418 TCGTAGCGCGGCCCRRTTGGGGCGTCCGAAGGGCTCTCGCACACGGGTGTAGGGTCTTGG 477
QY      122 AGGACGGGCTAACTATGCAACAGGAAATTTACCGGTTCCCGATCTCATGGGGTATATCCCGC 61
Db      538 TTGCTCTTCTCGTCTGACCGTTCCGGCTCTCGAGTTCCTACCGAAATGCCTCTG 597
QY      242 GATTTATCATGTTACCAATGATGCGCAAACTCTTCATAGTCTATAGGCGAGATAACC 301
Db      598 GATTTATCATGTTACCAATGATGCGCAAACTCTTCATAGTCTATAGGCGAGATAACC 657
QY      302 TGATCTCTACAGCACTGTTGCGTCCCTTGTGTCATGACAGTAATGTAGTAGTGTCT 361
Db      658 TGATCTCTACAGCACTGTTGCGTCCCTTGTGTCATGACAGTAATGTAGTAGTGTCT 717
QY      362 GGGTCCAAATTACCCCTACACTGTCTAGCCCCCGAGCCTCGAGAGCAGTCAAGGCTCTCTTTC 421

```

```

Db      718 GGGTCCAAATTACCCCTACACTGTGAGCCCCGAGCCTCGAGCAGTCAAGGCTCTCTTTC 777
QY      422 GGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTCTGTCTCGGTTTATAGTAGGAG 481
Db      778 GGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTCTGTCTCGGTTTATAGTAGGAG 837
QY      482 ACGCGTGTGGGCACTATTCTTTGGTAGGCCAAATGTTCACTATAGGCTCTGCCACGACG 541
Db      838 ACGCGTGTGGGCACTATTCTTTGGTAGGCCAAATGTTCACTATAGGCTCTGCCACGACG 897
QY      542 CTACGGTGCAGAACTGCAACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATGG 601
Db      898 CTACGGTGCAGAACTGCAACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATGG 957
QY      602 CA 603
Db      958 CA 959

RESULT 14
LOCUS   AX032171
DEFINITION
Sequence 53 from Patent EP0984067.
ACCESSION AX032171
VERSION   AX032171.1 GI:10279234
KEYWORDS
SOURCE   unidentified
ORGANISM unclassified.
REFERENCE
1. Maertens, G. and Stuyver, L.
Sequences of hepatitis c virus genotypes and their use as
therapeutic and diagnostic agents
Patent: EP 0984067-A 53 08-MAR-2000;
INNOGENETICS NV (BE)
FEATURES
source
Location/Qualifiers
1..959
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/clone="PC C/E1"
3..959
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:10279235"
/translation="MSTNPKPORKYNTNRRPQDKVFPGGQIVGGVLLPRRPRGM
GVRATRKTSERSQPRRRQPIPKARQPTGRSGWQPGYPMPLVYANGLGWAGLLSPRG
SRPNMGPNDPRKSRNLGKVIDLTLCGFADLMGYIPLVGGPXXGVARALAHGVRLSD
GVNYATGNLPGCSFSIFILALLSLCTVPASAVPYRNASGIYHVTNDPCNSSIYVEADN
LILHAPGCVPCVMTGNVSRWCVOITPTLSAPSLGAVTAPLRRAVDYLAGGAALCSALY
VGDACGALPLVQGMFTYRPRQATVQNCNSIYSGHVTGHRMA"
BASE COUNT 188 a 285 c 269 g 214 t 3 others
ORIGIN
Query Match 95.4%; Score 600.8; DB 6; Length 959;
Best Local Similarity 99.5%; Pred. No. 5.1e-162;
Matches 599; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 TGGGTAAGGTCTATCGATACCTAAACGTGCGGATTCGCCGATCTCATGGGGTATATCCCGC 61
Db      358 TGGGTAAGGTCTATCGATACCTAAACGTGCGGATTCGCCGATCTCATGGGGTATATCCCGC 417
QY      62 TCGTAGCGCGGCCCATTTGGGGCGTCCGAAGGGCTCTCGCACACGGGTGTAGGGTCTTGG 121
Db      418 TCGTAGCGCGGCCCRRTTGGGGCGTCCGAAGGGCTCTCGCACACGGGTGTAGGGTCTTGG 477
QY      122 AGGACGGGCTAACTATGCAACAGGAAATTTACCGGTTCCCGATCTCATGGGGTATATCCCGC 181
Db      478 AGGACGGGCTAACTATGCAACAGGAAATTTACCGGTTCCCGATCTCATGGGGTATATCCCGC 537
QY      182 TTGCTCTTCTCGTCTGACCGTTCCGGCTCTCGAGTTCCTACCGAAATGCCTCTG 241

```


Db	538	TTGCTCTTCTCGTGTCTGACCGTTCGGGCTCTGAGTTCCTACCGAAATGCCTCTG	597
Qy	242	GGATTATCATGTTACCAATGATGCCCCAAACTCTTCCATAGTCTATGAGCAGATAACC	301
Db	598	GGATTATCATGTTACCAATGATGCCCCAAACTCTTCCATAGTCTATGAGCAGATAACC	657
Qy	302	TGATCTACACGCACTGTTGGGTGCGCTTGCTGATGACAGGTAATGTGAGTAGATGCT	361
Db	658	TGATCTACACGCACTGTTGGGTGCGCTTGCTGATGACAGGTAATGTGAGTAGATGCT	717
Qy	362	GGGTCCAAATTTACCCCTACACTGTTCAGCCCCGAGCCTCGAGCAGTTCACGGCTCCTCTTC	421
Db	718	GGGTCCAAATTTACCCCTACACTGTTCAGCCCCGAGCCTCGAGCAGTTCACGGCTCCTCTTC	777
Qy	422	GGAGAGCGGTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCGGTTATACGTAGGAG	481
Db	778	GGAGAGCGGTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCGGTTATACGTAGGAG	837
Qy	482	ACCGGTGTGGGGCACTATTCTGTTAGGGCCAAATGTTCACTATAGGCTCGCCAGCAGC	541
Db	838	ACCGGTGTGGGGCACTATTCTGTTAGGGCCAAATGTTCACTATAGGCTCGCCAGCAGC	897
Qy	542	CTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGCGCATGTTACCGGCCACCGATGG	601
Db	898	CTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGCGCATGTTACCGGCCACCGATGG	957
Qy	602	CA 603	
Db	958	CA 959	

RESULT 15

BD172150

LOCUS

DEFINITION

New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy.

ACCESSION

BD172150

VERSION

BD172150.1

GI:28413448

KEYWORDS

JP 2002233389-A/27.

SOURCE

unidentified

ORGANISM

unclassified.

REFERENCE

1 (bases 1 to 959)

AUTHORS

Maertens, G. and Stuyver, L.

TITLE

New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy

JOURNAL

Patent: JP 2002233389-A 27 20-AUG-2002;

COMMENT

NV INNOGENETICS SA

OS

Unidentified

PN

JP 2002233389-A/27

PD

20-AUG-2002

PF

21-NOV-2001 JP 2001356707

PR

27-APR-1993 EP 93401099.2.05-AUG-1993 EP 93402019.9 PI

GEERT

MAERTENS, LIEVEN STUYVER

PC

C12N15/09, A61K35/76, A61K38/00, A61K39/00, A61K39/395, A61K39/395,

PC

A61K48/00,

PC

A61P31/20, C07K14/18, C07K16/10, C12Q1/68, G01N33/53, G01N33/53, PC

G01N33/566,

PC

G01N33/576, C12N15/00, A61K37/02

CC

Strandedness: Single;

CC

Topology: Linear;

CC

New sequences of hepatitis C virus genotypes for diagnosis,

CC

prophylaxis

CC

and therapy

CC

key

PH

key

FT

CDS

FEATURES

source

1. .959

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

188 a 269 g 214 t 3 others

BASE COUNT

188 a 269 g 214 t 3 others

ORIGIN	Query Match	95.4%;	Score 600.8;	DB 6;	Length 959;
	Best Local Similarity	99.5%;	Pred. No. S.le-162;	Indels	0;
	Matches	599;	Conservative	Mismatches	0;
Qy	2	TTGGTAAGGTCATCGATACCTAAACGTCGGGATTCGCCGATCTCATGGGGTATATCCCGC	61		
Db	358	TTGGTAAGGTCATCGATACCTAAACGTCGGGATTCGCCGATCTCATGGGGTATATCCCGC	417		
Qy	62	TCGTAGCGGCCCCCATTTGGGGCGTTCGCAAGGCTCTCCACACGCTGTGAGGGTCTCTTG	121		
Db	418	TCGTAGCGGCCCCCATTTGGGGCGTTCGCAAGGCTCTCCACACGCTGTGAGGGTCTCTTG	477		
Qy	122	AGGACGGGTAAACTATGCAACAGGGAATTTACCCGGTTCCTCTCTCTATCTTTATTC	181		
Db	478	AGGACGGGTAAACTATGCAACAGGGAATTTACCCGGTTCCTCTCTCTATCTTTATTC	537		
Qy	182	TTGCTCTTCTCTCGTGTCTGACCGTTCCGGCTCTGCGAGTTCCTTACCGAAATGCCCTCTG	241		
Db	538	TTGCTCTTCTCTCGTGTCTGACCGTTCCGGCTCTGCGAGTTCCTTACCGAAATGCCCTCTG	597		
Qy	242	GGATTTATCATGTTACCAATGATGCCCCAAACTCTTCCATAGTCTATGAGCAGATAACC	301		
Db	598	GGATTTATCATGTTACCAATGATGCCCCAAACTCTTCCATAGTCTATGAGCAGATAACC	657		
Qy	302	TGATCTACACGCACTGTTGGGTGCGCTTGCTGATGACAGGTAATGTGAGTAGATGCT	361		
Db	658	TGATCTACACGCACTGTTGGGTGCGCTTGCTGATGACAGGTAATGTGAGTAGATGCT	717		
Qy	362	GGGTCCAAATTTACCCCTACACTGTTCAGCCCCGAGCCTCGGAGCAGTCAAGGCTCTCTTC	421		
Db	718	GGGTCCAAATTTACCCCTACACTGTTCAGCCCCGAGCCTCGGAGCAGTCAAGGCTCTCTTC	777		
Qy	422	GGAGAGCGGTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCGGTTATACGTAGGAG	481		
Db	778	GGAGAGCGGTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCGGTTATACGTAGGAG	837		
Qy	482	ACCGGTGTGGGGCACTATTCTGTTAGGGCCAAATGTTCACTATAGGCTCGCCAGCAGC	541		
Db	838	ACCGGTGTGGGGCACTATTCTGTTAGGGCCAAATGTTCACTATAGGCTCGCCAGCAGC	897		
Qy	542	CTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGCGCATGTTACCGGCCACCGATGG	601		
Db	898	CTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGCGCATGTTACCGGCCACCGATGG	957		
Qy	602	CA 603			
Db	958	CA 959			

Search completed: December 20, 2003, 02:02:09
Job time : 2387.36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 174.169 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-31
Perfect score: 630
Sequence: 1 ATGGGTAGGTCATGCATAC.....TGATGATGACTGGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq 19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	100.0	630	17	AA12966 HCV El construct H
2	630	100.0	630	24	AA148930 Hepatitis C virus
3	602	95.6	959	15	AAQ78047 Hepatitis C virus
4	600.8	95.4	959	15	AAQ78049 Hepatitis C virus
5	597.2	94.8	959	15	AAQ78048 Hepatitis C virus
6	580	92.1	580	15	AAQ78045 Hepatitis C virus
7	580	92.1	580	15	AAQ78046 Hepatitis C virus
8	573.2	91.0	579	15	AAQ78092 Hepatitis C virus

9	510.2	81.0	579	15	AAQ78113 Hepatitis C virus
10	510.2	81.0	579	15	AAQ78093 Hepatitis C virus
11	347.2	55.1	576	16	AAQ83891 Hepatitis C virus
12	347.2	55.1	576	17	AA116605 Hepatitis C virus
13	344	54.6	576	16	AAQ83893 Hepatitis C virus
14	344	54.6	576	17	AA116607 Hepatitis C virus
15	340.8	54.1	576	16	AAQ83892 Hepatitis C virus
16	340.8	54.1	576	17	AA116606 Hepatitis C virus
17	337.6	53.6	576	16	AAQ83894 Hepatitis C virus
18	337.6	53.6	576	16	AAQ83890 Hepatitis C virus
19	337.6	53.6	576	17	AA116604 Hepatitis C virus
20	337.6	53.6	576	17	AA116608 Hepatitis C virus
21	335.6	53.3	630	17	AA112965 HCV El construct H
22	335.6	53.3	630	24	AA148929 Hepatitis C virus
23	331.6	52.6	9595	20	AA248443 Infectious hepatitis
24	331.6	52.6	9595	22	AA248443 Infectious hepatitis
25	331.6	52.6	9595	22	AA248443 Infectious hepatitis
26	331.6	52.6	9599	20	AA248443 Infectious hepatitis
27	331.2	52.6	576	16	AAQ83889 Hepatitis C virus
28	331.2	52.6	576	17	AA116603 Hepatitis C virus
29	326.8	51.9	2187	19	ABA03491 Cuticle protein 1
30	326.8	51.9	2540	14	AAQ43889 Nucleotide sequenc
31	326.8	51.9	2540	15	AAQ63753 Fragment of NANB h
32	325.2	51.6	1863	12	AAQ15363 NANB hepatitis vir
33	325.2	51.6	1880	13	AAQ24467 NANB hepatitis vir
34	325.2	51.6	2540	13	AAQ29628 Hepatitis C virus
35	324.8	51.6	795	17	AA112705 HCV El construct H
36	324.8	51.6	795	24	AA148914 Hepatitis C virus
37	324.8	51.6	2082	24	AA148939 Hepatitis C virus
38	324.8	51.6	2086	17	AA112973 HCV El construct H
39	324.8	51.6	2229	19	ABA03487 Cuticle protein 1
40	324.8	51.6	2433	17	AA112974 HCV El construct H
41	323.6	51.4	2116	12	AAQ12242 Encodes PT-NANBH v
42	323.6	51.4	2187	19	ABA03492 Cuticle protein 1
43	323.4	51.3	633	17	AA112706 HCV El construct H
44	323.4	51.3	633	24	AA148915 Hepatitis C virus
45	323.2	51.3	9472	14	AAQ33282 Korean hepatitis C

ALIGNMENTS

RESULT 1	
AA12966	
ID	AA12966 standard; DNA; 630 BP.
XX	AA12966;
AC	AA12966;
XX	24-SEP-1996 (first entry)
DT	HCV El construct HC163.
DE	HCV El construct HC163.
XX	HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera; ss.
XX	Hepatitis C virus.
OS	WO9604385-A2.
XX	15-FEB-1996.
PD	31-JUL-1995; 95WO-EP03031.
PF	29-JUL-1994; 94EP-0870132.
XX	(INNO-) INNOGENETICS NV.
XX	Bosman F, Buyse M, De Martynoff G, Maertens G;
PI	WPI; 1996-129401/13.
XX	Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT	

AAQ78049	DB	418	TCGTAGGCGGCCCR	TTGGGGCGTTCGCAAGGCTCTCCACACGGTGTGAGGGTCTTTG	477
ID AAQ78049 standard; cDNA; 959 BP.					
XX AC					
XX AAQ78049;					
XX					
25-MAR-2003 (updated)					
DT					
02-AUG-1995 (first entry)					
DT					
XX					
DE Hepatitis C virus E1 region.					
XX					
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;					
KW classification; immunisation; prophylaxis; serotyping; ss.					
KW					
XX					
OS Hepatitis C virus type 5a.					
XX					
XX Key Location/Qualifiers					
FT CDS 3..959					
FT /*tag= a					
FT /product= E1 polypeptide.					
XX					
XX WO9425601-A2.					
XX					
10-NOV-1994.					
XX					
27-APR-1994; 94WO-EP01323.					
XX					
27-APR-1993; 93EP-0401099.					
PR					
05-AUG-1993; 93EP-0402019.					
XX					
XX (INNO-) INNOGENETICS NV SA.					
PA					
XX					
XX Maertens G, Stuyver L;					
XX					
XX WPI: 1994-358277/44.					
DR					
DR P-PSDB; AAR63297.					
XX					
XX New polynucleotide sequences from hepatitis C virus - and related					
PT vectors, polypeptide(s) and antibodies, useful for immunisation,					
PT treatment, diagnosis and typing of HCV isolates					
XX					
XX Claim 3; Page 144; 404pp; English.					
XX					
XX Compositions comprising at least 5, and pref. 8 or more contiguous					
CC nucleotides selected from an HCV type 3 genomic sequence, more					
CC particularly (i) the region spanning positions 417-957 of the					
CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions					
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning					
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the					
CC region spanning positions 8023-8235 of the NS5 region of the BR36					
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic					
CC sequence, or, from a subtype 2d genomic sequence, a type 4 genomic					
CC sequence; or the coding region of subtype 5a, may be used as primers					
CC to amplify nucleic acid from an isolate belonging to a specific					
CC genotype, or as a probe for specific detection/classification of					
CC nucleic acid. Polypeptides encoded by the nucleotides in such					
CC compositions may be used for immunisation against HCV, for the					
CC detection of antibodies directed against HCV and for serotyping.					
CC This sequence corresponds to the E1 region of HCV subtype 5a and					
CC is taken from a clone designated PC C/E1.					
CC (Updated on 25-MAR-2003 to correct PN field.)					
XX					
SQ Sequence 959 BP; 188 A; 285 C; 269 G; 214 T; 3 other;					
Query Match 95.4%; Score 600.8; DB 15; Length 959;					
Best Local Similarity 99.5%; Pred. No. 4.6e-177;					
Matches 599; Conservative 3; Mismatches 0; Indels 0; Gaps 0;					
XX					
2 TGGGTAAGGTCATCGATACCCCTAACGTCGGATTCCCGATCTCATGCGGTATATCCGCG					61
XX					
358 TGGGTAAGGTCATCGATACCCCTAACGTCGGATTCCCGATCTCATGCGGTATATCCGCG					417
XX					
62 TCGTAGGCGGCCCATTTGGGGCGTTCGCAAGGCTCTCCACACGGTGTGAGGGTCTTTG					121
XX					
XX					

DR WPI: 1994-358277/44.
DR P-PSDB; AAR63296.
XX
PT New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates
XX
PS Claim 3; Page 141-142; 404pp; English.
XX
CC Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the
CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
CC region spanning positions 8023-8235 of the NS5 region of the BR36
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
CC sequence, or, from a subtype 2d genomic sequence, a type 4 genomic
CC sequence; or, the coding region of subtype 5a, may be used as primers
CC to amplify nucleic acid from an isolate belonging to a specific
CC genotype, or as a probe for specific detection/classification of
CC nucleic acid. Polypeptides encoded by the nucleotides in such
CC compositions may be used for immunisation against HCV, for the
CC detection of antibodies directed against HCV and for serotyping.
CC This sequence corresponds to the E1 region of HCV subtype 5a and
CC is taken from a clone designated PC-3-8.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 959 BP; 188 A; 287 C; 270 G; 214 T; 0 other;
Query Match 94.8%; Score 597.2; DB 15; Length 959;
Best Local Similarity 99.5%; Pred. No. 6.1e-176;
Matches 599; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 TGGGTAAGGTATCGATACCTTAACGTGCGGATTCGCGCATCTCATGGGGTATATCCCGC 61
DB 358 TGGGTAAGGTATCGATACCTTAACGTGCGGATTCGCGCATCTCATGGGGTATATCCCGC 417
QY 62 TCGTAGGGCGGCCCAATGTTGGGGCGTTCGCAAGGCTCTCGCACACGGTGTGAGGGTCTTGT 121
DB 418 TCGTAGGGCGGCCCGTGTGGGGCGTTCGCAAGGCTCTCGCACACGGTGTGAGGGTCTTGT 477
QY 122 AGGACGGGTAACTATGCAACAGGGAATTTACCGGTTGCTTTCTCTATCTTTATTC 181
DB 478 AGGACGGGTAACTATGCAACAGGGAATTTACCGGTTGCTTTCTCTATCTTTATTC 537
QY 182 TTGCTCTTCTCTGCTGCTGACCGTTCGCGGCTCTGCAAGTTCCTACCGAAATGCTCTG 241
DB 538 TTGCTCTTCTCTGCTGCTGACCGTTCGCGGCTCTGCAAGTTCCTACCGAAATGCTCTG 597
QY 242 GGATTTATCATGTTACCAATGATTGCCAAACTCTTCCATAGTCTATGAGGAGATAACC 301
DB 598 GGATTTATCATGTTACCAATGATTGCCAAACTCTTCCATAGTCTATGAGGAGATAACC 657
QY 302 TGATCTTACACGACCTGTTGGTGGCTTGTGTGTCATGACAGTAATGTAGTAGTGTCT 361
DB 658 TGATCTTACACGACCTGTTGGTGGCTTGTGTGTCATGACAGTAATGTAGTAGTGTCT 717
QY 362 GGGTCCAAATACCTTACACTGTACGCTCGAGCTCGAGCTCGAGCAGTCCCTCTTC 421
DB 718 GGGTCCAAATACCTTACACTGTACGCTCGAGCTCGAGCTCGAGCAGTCCCTCTTC 777
QY 422 GGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCGGGTTATAGTAGGAG 481
DB 778 GGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCGGGTTATAGTAGGAG 837
QY 482 ACGCGTGTGGGCACTATTCTTGGTAGGCCAAATGTTACCTATAGGCTTCGCCAGCAG 541
DB 838 ACGCGTGTGGGCACTATTCTTGGTAGGCCAAATGTTACCTATAGGCTTCGCCAGCAG 897
QY 542 CTACGGTGCAGAACTGCACTGTTTCAATTTACGTGGCCATGTTACCGGCCACCGATGG 601
DB 898 CTACGGTGCAGAACTGCACTGTTTCAATTTACGTGGCCATGTTACCGGCCACCGATGG 957

QY 602 CA 603
DB 958 CA 959
RESULT 6
AAQ78045
ID AAQ78045 standard; cDNA; 580 BP.
XX
AC AAQ78045;
XX
DT 25-MAR-2003 (updated)
DT 02-AUG-1995 (first entry)
XX
DE Hepatitis C virus E1 region.
XX
KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.
XX
OS Hepatitis C virus type 5a.
FH Key Location/Qualifiers
CDS 2..580
FT /*tag= a
FT /product= E1 polypeptide.
XX
PN WO9425601-A2.
XX
PD 10-NOV-1994.
XX
PF 27-APR-1994; 94WO-EP01323.
XX
PR 27-APR-1993; 93EP-0401099.
PR 05-AUG-1993; 93EP-0402019.
XX
PA (INNO-) INNOGENETICS NV SA.
XX
PI Maertens G, Stuyver L;
XX
DR WPI: 1994-358277/44.
DR P-PSDB; AAR63293.
XX
PT New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates
XX
PS Claim 3; Page 133-134; 404pp; English.
XX
CC Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the
CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
CC region spanning positions 8023-8235 of the NS5 region of the BR36
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
CC sequence, or, from a subtype 2d genomic sequence, a type 4 genomic
CC sequence; or, the coding region of subtype 5a, may be used as primers
CC to amplify nucleic acid from an isolate belonging to a specific
CC genotype, or as a probe for specific detection/classification of
CC nucleic acid. Polypeptides encoded by the nucleotides in such
CC compositions may be used for immunisation against HCV, for the
CC detection of antibodies directed against HCV and for serotyping.
CC This sequence corresponds to the E1 region of HCV subtype 5a and
CC is taken from a clone designated PC-4-1.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 580 BP; 111 A; 165 C; 155 G; 149 T; 0 other;
Query Match 92.1%; Score 580; DB 15; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.1e-170;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AACGTGCGGATTCGCGATCTCATGGGTATATCCGCTCGTAGCGGCCCATTTGGGG 83
DB 1 AACGTGCGGATTCGCGATCTCATGGGTATATCCGCTCGTAGCGGCCCATTTGGGG 60
QY 84 CGTGCAGAGGCTCTCGCACACGGGTGAGGGTCTTGAAGACGGGTAAACTATGCAAC 143
DB 61 CGTGCAGAGGCTCTCGCACACGGGTGAGGGTCTTGAAGACGGGTAAACTATGCAAC 120
QY 144 AGGGAATTTACCGGTTGCTCTTCTATCTATCTTATTTCTTCTCTCTCTCTCTGAC 203
DB 121 AGGGAATTTACCGGTTGCTCTTCTATCTTATCTTATCTTCTCTCTCTCTGAC 180
QY 204 CGTTCGGGCTCTGCAAGTTCCTTACCGAATGCTCTGGGATTTATCATGTTACCAATGA 263
DB 181 CGTTCGGGCTCTGCAAGTTCCTTACCGAATGCTCTGGGATTTATCATGTTACCAATGA 240
QY 264 TTGCCCCAAACTCTTCCATAGTCTATGAGGAGATTAACCTGATCCTACACGACCTGGTTG 323
DB 241 TTGCCCCAAACTCTTCCATAGTCTATGAGGAGATTAACCTGATCCTACACGACCTGGTTG 300
QY 324 CGTGCCTTGTATGATGACAGGTAAATGTAGTAGTATCTTCTATCTTCTCTCTCTGAC 383
DB 301 CGTGCCTTGTATGATGACAGGTAAATGTAGTAGTATCTTCTATCTTCTCTCTGAC 360
QY 384 GTACAGCCCGAGCTCTCGGACGAGTACGGTCTCTTTCGAGAGCGGTTGACTACCTAGC 443
DB 361 GTACAGCCCGAGCTCTCGGACGAGTACGGTCTCTTTCGAGAGCGGTTGACTACCTAGC 420
QY 444 GGGAGGGGCTCGCCTCTGCTCCGGTTATAGTAGGAGCGCTGTGGGCACTATTCTT 503
DB 421 GGGAGGGGCTCGCCTCTGCTCCGGTTATAGTAGGAGCGCTGTGGGCACTATTCTT 480
QY 504 GGTAGGCGCAATGTTTACCTATAGGCTCGCCAGCAGCTACGGTGCAGAACTGCAACTG 563
DB 481 GGTAGGCGCAATGTTTACCTATAGGCTCGCCAGCAGCTACGGTGCAGAACTGCAACTG 540
QY 564 TTCCATTTACAGTGGCCATGTTACCGGCCACCGGATGGCA 603
DB 541 TTCCATTTACAGTGGCCATGTTACCGGCCACCGGATGGCA 580

RESULT 7

AAQ78046
ID AAQ78046 standard; cDNA; 580 BP.
XX AC AAQ78046;
XX XX
XX 25-MAR-2003 (updated)
DT 02-AUG-1995 (first entry)
XX DE
XX Hepatitis C virus E1 region.
XX KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
XX KW Classification; immunisation; prophylaxis; serotyping; ss.
XX OS
XX Hepatitis C virus type 5a.
XX FH Key Location/Qualifiers
XX CDS 2..580
FT /*tag= a
FT /product= E1 polypeptide.
XX PN W09425601-A2.
XX PD 10-NOV-1994.
XX XX 27-APR-1994; 94WO-EP01323.
XX PF
XX 27-APR-1993; 93BP-0401099.
XX PR 05-AUG-1993; 93BP-0402019.
XX XX
XX (INNO-) INNOGENETICS NV SA.

XX FI Maertens G, Stuyver L;
XX DR WPI; 1994-358277/44.
XX DR P-PSDB; AAR63294.
XX New polynucleotide sequences from hepatitis C virus - and related
XX vectors, polypeptide(s) and antibodies, useful for immunisation,
XX treatment, diagnosis and typing of HCV isolates
XX
XX Claim 3; Page 135-136; 404pp; English.
XX
XX Compositions comprising at least 5, and pref. 8 or more contiguous
XX nucleotides selected from an HCV type 3 genomic sequence, more
XX particularly (i) the region spanning positions 417-957 of the
XX Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
XX 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
XX positions 4892-5292 of the NS3/4 region of the NS5 region of the BR36
XX region spanning positions 8023-8235 of the NS5 region of the BR36
XX subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
XX sequence, or from a subtype 2d genomic sequence, a type 4 genomic
XX sequence; or the coding region of subtype 5a, may be used as primers
XX to amplify nucleic acid from an isolate belonging to a specific
XX genotype, or as a probe for specific detection/classification of
XX nucleic acid. Polypeptides encoded by the nucleotides in such
XX compositions may be used for immunisation against HCV, for the
XX detection of antibodies directed against HCV and for serotyping.
XX This sequence corresponds to the E1 region of HCV subtype 5a and
XX is taken from a clone designated PC-4-6.
XX (Updated on 25-MAR-2003 to correct PN field.)

Sequence 580 BP; 111 A; 165 C; 155 G; 149 T; 0 other;

Query Match 92.1%; Score 580; DB 15; Length 580;

Best Local Similarity 100.0%; Pred. No. 1.1e-170;

Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AACGTGCGGATTCGCGATCTCATGGGTATATCCGCTCGTAGCGGCCCATTTGGGG 83
DB 1 AACGTGCGGATTCGCGATCTCATGGGTATATCCGCTCGTAGCGGCCCATTTGGGG 60
QY 84 CGTGCAGAGGCTCTCGCACACGGGTGAGGGTCTTGAAGACGGGTAAACTATGCAAC 143
DB 61 CGTGCAGAGGCTCTCGCACACGGGTGAGGGTCTTGAAGACGGGTAAACTATGCAAC 120
QY 144 AGGGAATTTACCGGTTGCTCTTCTATCTTATCTTCTCTCTCTCTCTGAC 203
DB 121 AGGGAATTTACCGGTTGCTCTTCTATCTTATCTTCTCTCTCTCTGAC 180
QY 204 CGTTCGGGCTCTGCAAGTTCCTTACCGAATGCTCTGGGATTTATCATGTTACCAATGA 263
DB 181 CGTTCGGGCTCTGCAAGTTCCTTACCGAATGCTCTGGGATTTATCATGTTACCAATGA 240
QY 264 TTGCCCCAAACTCTTCCATAGTCTATGAGGAGATTAACCTGATCCTACACGACCTGGTTG 323
DB 241 TTGCCCCAAACTCTTCCATAGTCTATGAGGAGATTAACCTGATCCTACACGACCTGGTTG 300
QY 324 CGTGCCTTGTATGATGACAGGTAAATGTAGTAGTATCTTCTATCTTCTCTCTGAC 383
DB 301 CGTGCCTTGTATGATGACAGGTAAATGTAGTAGTATCTTCTATCTTCTCTGAC 360
QY 384 GTACAGCCCGAGCTCTCGGACGAGTACGGTCTCTTTCGAGAGCGGTTGACTACCTAGC 443
DB 361 GTACAGCCCGAGCTCTCGGACGAGTACGGTCTCTTTCGAGAGCGGTTGACTACCTAGC 420
QY 444 GGGAGGGGCTCGCCTCTGCTCCGGTTATAGTAGGAGCGCTGTGGGCACTATTCTT 503
DB 421 GGGAGGGGCTCGCCTCTGCTCCGGTTATAGTAGGAGCGCTGTGGGCACTATTCTT 480
QY 504 GGTAGGCGCAATGTTTACCTATAGGCTCGCCAGCAGCTACGGTGCAGAACTGCAACTG 563
DB 481 GGTAGGCGCAATGTTTACCTATAGGCTCGCCAGCAGCTACGGTGCAGAACTGCAACTG 540

QY 564 TTCCATTTACAGTGGCCCATGTTACCGGCCACCGGATGGCA 603
 Db |||||
 541 TTCCATTTACAGTGGCCCATGTTACCGGCCACCGGATGGCA 580

RESULT 8

AAQ78092
 ID AAQ78092 standard; cDNA; 579 BP.

AC AAQ78092;

DT 25-MAR-2003 (updated)

DT 15-AUG-1995 (first entry)

DE Hepatitis C virus E1 region.

KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping; ss.

OS Hepatitis C virus.

FH Key Location/Qualifiers

FT CDS 1..579

FT /*tag= a

FT /product= E1 polypeptide.

FT mat_peptide 1..576

FT /*tag= b

XX WO9425601-A2.

PN 10-NOV-1994.

XX 27-APR-1994; 94WO-EP01323.

XX 27-APR-1993; 93EP-0401099.

PR 05-AUG-1993; 93EP-0402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

PI WPI: 1994-358277/44.

DR P-PSDB; AAR63354.

XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates

PS Claim 3; Page 202-203; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the
 CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
 CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
 CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
 CC region spanning positions 8023-8235 of the NS5 region of the BR36
 CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
 CC sequence, may be used as primers to amplify nucleic acid from an
 CC isolate belonging to a specific genotype, or as a probe for specific
 CC detection/classification of nucleic acid. Polypeptides encoded by
 CC the nucleotides in such compositions may be used for immunisation
 CC against HCV, for the detection of antibodies directed against HCV
 CC and for serotyping. This sequence corresponds to the E1 region
 CC of HCV.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 579 BP; 108 A; 167 C; 156 G; 148 T; 0 other;

Query Match 91.0%; Score 573.2; DB 15; Length 579;
 Best Local Similarity 99.5%; Pred. No. 1.5e-168;
 Matches 575; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 25 ACGTGGGATTCCCGGATCTCATGGGGTATATCCCGCTCGTAGGGCGGCCCAATTGGGGGC 84
 Db |||||
 1 ACGTGGGATTCCCGGATCTCATGGGGTACATCCCGCTCGTAGGGCGGCCCGTTGGGGGC 60
 QY 85 GTCGCAAGGGCTCTCGCACACGGTGTGAGGTCCTTGAGAGCGGGTAACTATGCAACA 144
 Db |||||
 61 GTCGCAAGGGCTCTCGCACACGGTGTGAGGTCCTTGAGAGCGGGTAACTATGCAACA 120
 QY 145 GGGAAATTTACCGGTTGCTCTTTCTATCTTTATTTCTTGTCTTCTCTCGTGTCTGACC 204
 Db |||||
 121 GGGAAATTTACCGGTTGCTCTTTCTATCTTTATTTCTTGTCTTCTCTCGTGTCTGACC 180
 QY 205 GTTCGGGCTCTGCGAGTTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGAT 264
 Db |||||
 181 GTTCGGGCTCTGCGAGTTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGAT 240
 QY 265 TCCCAAACTCTTCCATAGTCTATGAGGCAGATAACCTGATCTACACGCACCTGGTTGC 324
 Db |||||
 241 TCCCAAACTCTTCCATAGTCTATGAGGCAGATAACCTGATCTACACGCACCTGGTTGC 300
 QY 325 GTGCCTTGTTCATGACAGGTAATGTGAGTAGATGCTGGGTCCAAATTACCCCTACACTG 384
 Db |||||
 301 GTGCCTTGTTCATGACAGGTAATGTGAGTAGATGCTGGGTCCAAATTACCCCTACACTG 360
 QY 385 TCAGCCCCGAGCTCGGAGCAGTCACGGCTCTCTTCGGAGAGCGCTTGACTACTAGCG 444
 Db |||||
 361 TCAGCCCCGAGCTCGGAGCAGTCACGGCTCTCTTCGGAGAGCGCTTGACTACTAGCG 420
 QY 445 GGAGGGGCTGCCCTCTGTCTCCGTTATACGTAGGAGAGCGGTGTGGGCACTATTCTTG 504
 Db |||||
 421 GGAGGGGCTGCCCTCTGTCTCCGTTATACGTAGGAGAGCGGTGTGGGCACTATTCTTG 480
 QY 505 GTAGGCCAAATGTTTACCTATAGGCTTCGCCAGCAGCTACGGTGCAGAACTGCAACTGT 564
 Db |||||
 481 GTAGGCCAAATGTTTACCTATAGGCTTCGCCAGCAGCTACGGTGCAGAACTGCAACTGT 540
 QY 565 TCATTTACAGTGGCCATGTTACCGGCCACCGGATGGC 602
 Db |||||
 541 TCCATTTACAGTGGCCATGTTACCGGCCACCGGATGGC 578

RESULT 9

AAQ78113

ID AAQ78113 standard; cDNA; 579 BP.

AC AAQ78113;

DT 25-MAR-2003 (updated)

DT 18-AUG-1995 (first entry)

XX Hepatitis C virus NS5B region.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping; ss.

XX Hepatitis C virus.

FH Key Location/Qualifiers

FT CDS 1..579

FT /*tag= a

FT /product= NS5B polypeptide.

FT mat_peptide 1..576

FT /*tag= b

XX WO9425601-A2.

XX 10-NOV-1994.

XX 27-APR-1994; 94WO-EP01323.

XX 27-APR-1993; 93EP-0401099.

XX 05-AUG-1993; 93EP-0402019.


```
PD 22-FEB-1996.
XX 15-AUG-1995; 95WO-US10398.
XX 15-AUG-1994; 94US-0290665.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (USSH ) US SEC DEPT HEALTH.
XX Bukh J, Miller RH, Purcell RH;
XX WPI; 1996-139709/14.
XX P-PSDB; AAR89551.
XX DNA and amino acid sequence of HCV envelope 1 and core proteins -
PT used to determine HCV genotype and as vaccines against HCV infection
XX Claim 1; Page 107; 340pp; English.
XX AAT16559-T16609 are cDNAs encoding the E1 (envelope-1) protein of 51 HCV
CC isolates. The isolated sequences are useful for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers
CC are also useful for HCV genotyping. Proteins encoded by the cDNAs
CC can be used in vaccines for immunising against HCV infection. The
CC proteins may also be used to detect antibodies against HCV in serum,
CC saliva, lymphocytes or other mononuclear cells. The antibodies may be
CC used in the prevention of HCV infection.
XX Sequence 576 BP; 106 A; 161 C; 168 G; 141 T; 0 other;
SQ
Query Match 55.1%; Score 347.2; DB 17; Length 576;
Best Local Similarity 90.7%; Pred. No. 5.3e-98;
Matches 370; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 220 GTTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 279
DB 1 GTCCCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 60
QY 280 ATAGTCTATGAGGAGATGAACCTGATCTTACACGACCTGTTGCGTGTGTCATG 339
DB 61 ATAGTCTATGAGGAGCTGACAACTGATCTTCCGGGTTTATCATGTTACCAATGATGCCCAAACTCTTCC 120
QY 340 ACAGTAAATGTCAGTAGATGCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 399
DB 121 GAAGCGGTTCACGGCTCTCTTCGGAGGCTGTTGACTACTTACGCGGAGGGCTGCGCTC 180
QY 400 GGAGCAGTCACGGCTCTCTTCGGAGGCTGTTGACTACTTACGCGGAGGGCTGCGCTC 459
DB 181 GGAGCGGTTCACGGCTCTCTTCGGAGGCTGTTGACTACTTACGCGGAGGGCTGCGCTC 240
QY 460 TGCTCCGGCTTATACGTAGGAGACCGCGTGTGGGGCACTTCTTGGTAGGCCAATGTC 519
DB 241 TGCTCCGGCTTATACGTAGGAGACCGCGTGTGGGGCACTTCTTGGTAGGCCAATGTC 300
QY 520 ACCTATAGCCCTCGCCAGCACCTACGCTGCGAGACTGCAACTGTTCCATTTACAGTGGC 579
DB 301 ACCTATAGCCCTCGCCAGCATACTAGTGTGAGGACTGCACTGTTCCATTTACAGGCGC 360
QY 580 CATGTTACCGGCGACCGGATGGCATGGGATGATGATGAACTGGTAA 627
DB 361 CATATCACCGGCGACCGAATGGCATGGGACATGATGATGAACTGGTCA 408
RESULT 13
ID AAQ83893
XX AAQ83893 standard; cDNA; 576 BP.
AC AAQ83893;
XX 25-MAR-2003 (updated)
DT 19-SEP-1995 (first entry)
XX Hepatitis C virus envelope 1 gene cDNA isolate SA7.
```

```
XX Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate SA7;
KW diagnosis; vaccines; antibodies; antisera; gene inhibition; sa.
XX Hepatitis C virus.
XX Key Location/Qualifiers
FT mat_peptide 1..576
FT /*tag= a
XX WO9501442-A2.
XX 12-JAN-1995.
XX 28-JUN-1994; 94WO-US07320.
XX 29-JUN-1993; 93US-0086428.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Bukh J, Miller RH, Purcell RH;
XX WPI; 1995-061006/08.
XX P-PSDB; AAR69682.
XX Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
PT and in vaccines
XX Claim 1; Page 78; 186pp; English.
XX AAQ83893 encodes AAR69682 hepatitis C virus (HCV) envelope 1 (E1)
CC protein isolate SA7, both can be used for the diagnosis of HCV
CC infection, and in the prodn. of anti-HCV vaccines, antibodies
CC and antisera. The cDNA may also be used to inhibit the expression
CC of the HCV E1 gene.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 576 BP; 104 A; 165 C; 168 G; 139 T; 0 other;
SQ
Query Match 54.6%; Score 344; DB 16; Length 576;
Best Local Similarity 90.2%; Pred. No. 5.3e-97;
Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 220 GTTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 279
DB 1 GTCCCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 60
QY 280 ATAGTCTATGAGGAGATGAACCTGATCTTACACGACCTGTTGCGTGTGTCATG 339
DB 61 ATAGTCTATGAGGCTGACAACTGATCTTCCGGGTTTATCATGTTACCAATGATGCCCAAACTCTTCC 120
QY 340 ACAGGTAATGTCAGTAGATGCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 399
DB 121 CAAATAATGTCAGTAGGCTGCTGGGTTCCAAATCACCCCCACATTTGTACGCCGACCTC 180
QY 400 GGAGCAGTCACGGCTCTCTTTCGGAGGCTGTTGACTACTTACGCGGAGGGCTGCGCTC 459
DB 181 GGAGCGGTTCACGGCTCTCTTTCGGAGGCTGTTGACTACTTACGCGGAGGGCTGCGCTC 240
QY 460 TGCTCCGGCTTATACGTAGGAGACCGCGTGTGGGGCACTTCTTGGTAGGCCAATGTC 519
DB 241 TGCTCCGGCTTATACGTAGGAGACCGCGTGTGGGGCACTTCTTGGTAGGCCAATGTC 300
QY 520 ACCTATAGCCCTCGCCAGCACCTACGCTGCGAGACTGCAACTGTTCCATTTACAGTGGC 579
DB 301 AGCTATAGCCCTCGCCAGCACACTACGCTGCGAGACTGCAACTGTTCCATTTACAGTGGC 360
QY 580 CATGTTACCGGCGACCGGATGGCATGGGATGATGATGAACTGGTAA 627
DB 361 CATATCACCGGCGACCGAATGGCATGGGACATGATGATGAACTGGTCA 408
```



```
RESULT 14
AAT16607
ID AAT16607 standard; cDNA; 576 BP.
XX
AC AAT16607;
XX
30-SEP-1996 (first entry)
XX
DE Hepatitis C virus isolate SA7 envelope 1 gene.
XX
KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 1..576
FT /tag= a
FT /product= envelope-1_protein
FT /note= "does not contain start or stop codon"
XX
PN W09605315-A2.
XX
22-FEB-1996.
XX
15-AUG-1995; 95WO-US10398.
XX
15-AUG-1994; 94US-0290665.
XX
(USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (USSH ) US SEC DEPT HEALTH.
XX
PI Bukh J, Miller RH, Purcell RH;
XX
WPI; 1996-139709/14.
DR P-PSDB; AAR89553.
XX
DNA and amino acid sequence of HCV envelope 1 and core proteins -
PT used to determine HCV genotype and as vaccines against HCV infection
XX
Claim 1; Page 108; 340pp; English.
XX
AAT16559-T16609 are cDNAs encoding the E1 (envelope-1) protein of 51 HCV
CC isolates. The isolated sequences are useful for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers
CC are also useful for HCV genotyping. Proteins encoded by the cDNAs
CC can be used in vaccines for immunising against HCV infection. The
CC proteins may also be used to detect antibodies against HCV in serum,
CC saliva, lymphocytes or other mononuclear cells. The antibodies may be
CC used in the prevention of HCV infection.
XX
SQ Sequence 576 BP; 104 A; 165 C; 168 G; 139 T; 0 other;

Query Match 54.6%; Score 344; DB 17; Length 576;
Best Local Similarity 90.2%; Pred. No. 5.3e-97;
Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAATGCCTCTGGGATTATCATGTACCAATGATGCGCCCAAACTCTTCC 279
DB 1 GTCCCTACCGAATGCCTCTGGGATTATCATGTACCAATGATGCGCCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGCAGATAACCTGTATCTACAGCACCTGGTTGCGTTCCTTGTGTATG 339
DB 61 ATAGTCTATGAGGCTGACAACTGTATCTGTACGACACCTGGTTGCGTTCCTGTGTATG 120

QY 340 ACAGGTAATGTAGTAGAGTCTGGTCCAAATACCCCTACACTGTACGCCCCGAGCCTC 399
DB 121 CAAAATAATGTAGTAGTGTCTGGGTCCAAATACCCCTACACTGTGTACGCCCCGAGCCTC 180

QY 400 GGAGCAGTACGCGCTCCTCTTCGGAGAGCGGTTGACTACTAGCGGAGGGGCTGCCCTC 459
DB 181 GGAGCGGTACGCGCTCCTCTTCGGAGAGCGGTTGACTACTAGCGGAGGGGCTGCCCTC 240

RESULT 15
AAQ83892
ID AAQ83892 standard; cDNA; 576 BP.
XX
AC AAQ83892;
XX
25-MAR-2003 (updated)
DT 19-SEP-1995 (first entry)
XX
DE Hepatitis C virus envelope 1 gene cDNA isolate SA6.
XX
KW Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate SA6;
KW diagnosis; vaccines; antibodies; antisera; gene inhibition; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT mat_peptide 1..576
FT /tag= a
XX
W09501442-A2.
XX
12-JAN-1995.
XX
28-JUN-1994; 94WO-US07320.
XX
29-JUN-1993; 93US-0086428.
XX
(USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Bukh J, Miller RH, Purcell RH;
XX
WPI; 1995-061006/08.
DR P-PSDB; AAR69681.
XX
Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
PT and in vaccines
XX
Claim 1; Pages 77-78; 186pp; English.
XX
AAQ83892 encodes AAR69681 hepatitis C virus (HCV) envelope 1 (E1)
CC protein isolate SA6, both can be used for the diagnosis of HCV
CC infection, and in the prodn. of anti-HCV vaccines, antibodies
CC and antisera. The cDNA may also be used to inhibit the expression
CC of the HCV E1 gene.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 576 BP; 101 A; 156 C; 171 G; 148 T; 0 other;

Query Match 54.1%; Score 340.8; DB 16; Length 576;
Best Local Similarity 89.7%; Pred. No. 5.2e-96;
Matches 366; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAATGCCTCTGGGATTATCATGTACCAATGATGCGCCCAAACTCTTCC 279
DB 1 GTTCCTACCGAATGCCTCTGGGATTATCATGTACCAATGATGCGCCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGCAGATAACCTGTATCTACAGCACCTGGTTGCGTTCCTTGTGTATG 339
DB 61 ATAGTCTATGAGGCTGACAACTGTATCTGTACGACACCTGGTTGCGTTCCTGTGTATG 120
```


Db	61	ATAGTCTATGAGGCTGATGACCTGATCCTTACACGACACCTGGCTGCGTCCCTGTGTCCGG	120
Qy	340	ACAGGTAATGTAGTAGATGCTGGGTCCAAATTACCCCTACACTGTCTAGCCCCGAGCCTC	399
Db	121	AAGGATAATGTCAGTAGATGCTGGGTTTATATCACCCCCACACTATCAGCCCCGAGCCTC	180
Qy	400	GGAGCAGTCAAGGCTCCTCTTCGGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTC	459
Db	181	GGAGCGGTACGGCTCCTCTTCGGAGGGCGTTGATTACTTGGCGGAGGGGCGGCCCTG	240
Qy	460	TGCTCCGCGTTATACGTAGAGACGCGTGTGGGGCACTATTCTTGTAGGCCCAATGTTTC	519
Db	241	TGCTCCGCGTTATACGTTCGGAGACGTTGTGGGGCATTTGTTTGTAGGCCAAATGTTTC	300
Qy	520	ACCTATAGGCTCGCCAGCAGCTACGCTGCAGAACTGCAACTGTTTCCATTTACAGTGGC	579
Db	301	ACCTATAGGCTCGCCAGCATGCTACGGTACAGGACTGCACCTGCTCCCATTTACAGTGGC	360
Qy	580	CATGTTACCGGCCACCGGATGGCATGGATATGATGAACCTGGTAA	627
Db	361	CATATCACTGGCCACCGGATGGCATGGACATGATGAATTGGTCA	408

Search completed: December 19, 2003, 18:51:24
Job time : 175.169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1620.71 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303a-31
Perfect score: 630
Sequence: 1 ATGGTAGGTCATCGATAC.....TCGATGACTGGTAATAG 630

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gsa1:*

29: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	57.6	9.1	488	9	AV755731
C 2	39.6	6.3	492	9	AV758366
C 3	38.4	6.1	920	29	CNS0062R
C 4	37.8	6.0	1002	13	BX374637

C 5	37	5.9	944	13	BX437004
C 6	37	5.9	1201	13	BX443048
C 7	36.6	5.8	703	10	BF479339
C 8	36.6	5.8	721	12	BI462364
C 9	36.4	5.8	962	29	CNS0140F
C 10	36.4	5.8	1201	13	BX356664
C 11	35.8	5.7	918	13	BX384351
C 12	35.4	5.6	405	9	AA021112
C 13	35.4	5.6	962	10	BG571230
C 14	35.4	5.6	1178	13	BX397483
C 15	35.2	5.6	727	13	BU425620
C 16	35.2	5.6	791	29	CNS00AKB
C 17	35	5.6	632	10	BG297765
C 18	35	5.6	903	28	AYS31723
C 19	34.8	5.5	1101	29	CNS002BV
C 20	34.6	5.5	715	29	AG031744
C 21	34.6	5.5	879	10	BG198380
C 22	34.4	5.5	448	12	BI478979
C 23	34.4	5.5	678	12	BM332179
C 24	34.4	5.5	771	12	BM348565
C 25	34.4	5.5	947	12	BI838659
C 26	34.2	5.4	229	10	AW932968
C 27	34.2	5.4	359	13	BY370241
C 28	34.2	5.4	458	28	BH121694
C 29	34.2	5.4	534	12	BJ082331
C 30	34.2	5.4	761	29	CNS020DP
C 31	34.2	5.4	1201	13	BX335532
C 32	34.2	5.4	1201	13	BX376097
C 33	34	5.4	560	9	AA607428
C 34	34	5.4	733	14	CD445238
C 35	33.8	5.4	690	29	CC315210
C 36	33.8	5.4	884	29	CNS02Y9J
C 37	33.8	5.4	1201	13	BX387694
C 38	33.6	5.3	331	10	BE040504
C 39	33.6	5.3	449	14	CA500407
C 40	33.6	5.3	682	28	BH316210
C 41	33.6	5.3	784	12	BG920015
C 42	33.6	5.3	847	14	CA987143
C 43	33.6	5.3	903	12	BG855135
C 44	33.6	5.3	1101	29	CNS002SC
C 45	33.6	5.3	1176	13	BU184099

ALIGNMENTS

RESULT 1	AV755731/c	488 bp	mRNA	linear	EST 19-OCT-2000
LOCUS	AV755731	BM Homo sapiens	CDNA clone BMFAKB03 5'		mRNA sequence.
DEFINITION	AV755731				
ACCESSION	AV755731				
VERSION	AV755731.1	GI:10913579			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 488)				
AUTHORS	Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.				
TITLE	Homo sapiens CDNA BM clones				
JOURNAL	Unpublished				
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex. 45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai. Location/Qualifiers				


```

source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN

Query Match 9.1%; Score 57.6; DB 9; Length 488;
Best Local Similarity 66.3%; Pred. No. 9.4e-05;
Matches 114; Conservative 0; Mismatches 54; Indels 4; Gaps 2;

QY 455 CCTCTCTCCGGTTATACGTAGGAGACCGGTGTGGGACATTTCTTGGTAGGCCAAA 514
Db 459 CGCTCTCTCAGCTCTCTACGTGTGGGACCTCTGCGACGGAGTGATGTCAGTTTCAGC 400

QY 515 TGTTCACCTATAGCTCGCAGCAGCAGCTACGCTGAGAGTCACTGTTTCCATTTACA 574
Db 399 TGATCATCT---GGCCTCAGCAGCAGCAGTGTGTGCGATGATGCACTGCTCCATCTATC 343

QY 575 GTGGCCATGTTTACCGGCCACCGGATG-GCATGGGATATGATGATGAAGTGGT 625
Db 342 CTGGCGCCATCTGGACACCGTATGAGCATGGGACATGATGATGAAGTGGT 291

RESULT 2
AV758366/c 492 bp mRNA linear EST 19-OCT-2000
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
Gu,Y., Li,N., Qian,B., Liu,P., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1. .492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN

Query Match 6.3%; Score 39.6; DB 9; Length 492;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 102; Conservative 0; Mismatches 64; Indels 4; Gaps 2;

QY 457 CTCTGTCCCGTTATACGTAGGAGACCGGTGTGGGACATTTCTTGTAGGCCAAATG 516
Db 458 CTGTGATCAGCTCACTACGTGTGGGACCTCTGCTGTGGGATTCGTTGACGCCCACTG 399

QY 517 TTCACTTATAGGCTCGCCAGCAGCTAGCGGTGAGCACTGTCACATTTTACAGT 576
Db 398 ATTATCTCT---CAGCAGCAACATTTGTTGTGCAAGAATGCAACTGCTCATTTCTATCCT 342

QY 577 GGCCATGTTTACCGG-CCACCGGATGCGATGGGATATGATGATGAAGTGGT 625
Db 341 GGCTGCATCACTGCACTACAGTATGCGATAGGCTATGATGATGAAGTGGT 292

CNS0062R 920 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR13J24 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL061710 GI:4943910
GSS.
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 920)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
Bp 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1. .920
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR13J24"
/clone_lib="RPCI-98"
/notes="end : TET3"
BASE COUNT 307 a 62 c 39 g 122 t 390 others
ORIGIN

Query Match 6.1%; Score 38.4; DB 29; Length 920;
Best Local Similarity 18.0%; Pred. No. 15;
Matches 55; Conservative 124; Mismatches 126; Indels 0; Gaps 0;

QY 93 GGCTCTCGCACCGTGTAGGCTCTTGGAGCGGGTAAATATGCAACAGGGAATT 152
Db 416 KBBKKKCAANBDBCYRKRSTGAVYGBKRGKGBKGYKXKSGSRDGAKBDAAGH 475

QY 153 ACCCGGTGTCTTCTCTATCTTTATTTCTTCTTCTTCTTCTGTCGTCGTCGGC 212
Db 476 DKRSTCYAVCKENANABYSBCTSKYBTSYGVAAATTTAAATTTKYAATCNCYMC 535

QY 213 CTCTGCACTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATTTGCCCAA 272

```


ACCESSION	BX443048
VERSION	EST.
KEYWORDS	Genome
SOURCE	Homo sapiens
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D. Full-length cDNA libraries and normalization Unpublished Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3085.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CSODG003AH04QP1&cluster=3085.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Paradise Avenue Genoscope sequence ID : CSODG003AH04QP1.
FEATURES	Location/Qualifiers source 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODG003Y007" /tissue_type="B CELLS (RAMOS CELL LINE)" /cell_line="RAMOS CELL LINE" /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)" /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized." BASE COUNT 349 a 184 c 261 g 350 t 57 others
ORIGIN	
Query Match	5.9%; Score 37; DB 13; Length 1201;
Best Local Similarity	41.1%; Pred.No. 37; Mismatches 108; Indels 0; Gaps 0;
Matches	90; Conservative
QY	140 CAACAGGAATTTACC GGCTGTCTTC TCTA CTATTATTTATCTCTCTCTCGTGTC 199
Db	1097 BMAAAAAAAWTTWACMKKGSYYTTWAATTTKTWTATATTTCA TTKTTATAAAGAKTCY 1038
QY	200 TGACGGTCCGCTCTGCAGTTCCTCCACGGAATGCCCTCTGGATTATCATGTATACCA 259
Db	1037 TMAAATGACTGCTGKAAGTBCTCTTTACAATGGCGTGGTGATGATCAGATTAWAA 978
QY	260 ATGATTGCCAAAATCTTCCATAGTCTATGAGGCAGATAACCTGATCTTACAGCACCTG 319
Db	977 ACTKTAACACCAAKTACTTTATGKAATTA KAAAKCTAACATTTGTATCTCAAACTTA 918
QY	320 GTTGGCTGCTTGTCATGACAGTATGTGAGTAGAT 358
Db	917 CTTGAATCAGTTTTCAATTCATTCATCAAATTAAGTTKAT 879
RESULT 7	
BF479339	
LOCUS	
DEFINITION	L48-3008T3 Ice plant Lambda Uni-zap XR expression library, 48 hours NaCl treatment Mesembryanthemum crystallinum CDNA clone L48-3008 5', mRNA sequence.
ACCESSION	BF479339
VERSION	EST.
KEYWORDS	Genome
SOURCE	Mesembryanthemum crystallinum (common iceplant)
ORGANISM	Vicia-Lentaceae; Streptophyta; Embryophyta; Tracheophyta;
REFERENCE	1 (bases 1 to 703) Cushman, J.C. An expressed sequence tag database for the common ice plant, Mesembryanthemum crystallinum Unpublished Contact: Cushman JC Department of Biochemistry University of Nevada MS200, Reno, NV 89557-0014, USA Tel: 775-784-1918 Fax: 775-784-1650 Email: jcushman@unr.edu PCR Primers FORWARD: T7 BACKWARD: T3 Plate: L48-31 row: A column: 8 Seq primer: T3 High quality sequence stop: 350 POLYA=No.
FEATURES	Location/Qualifiers source 1..703 /organism="Mesembryanthemum crystallinum" /mol_type="mRNA" /db_xref="taxon:3544" /clone="L48-3008" /tissue_type="Leaf, 48 h 0.4M NaCl" /dev_stage="Six week old" /clone_lib="Ice plant Lambda Uni-Zap XR expression library 48 hours NaCl treatment" /note="vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	170 a 150 c 178 g 205 t
ORIGIN	
Query Match	5.8%; Score 36.6; DB 10; Length 703;
Best Local Similarity	51.5%; Pred.No. 43;
Matches	84; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY	85 GTCCGAAGGCTCTCGCACGGGTGTCAGGGTCTCTTGAGGACGGGTAAACTATGCAACA 144
Db	529 GTGGAGTGGGGTTTCCACCTCGTATGATGATGGTGAGGAGCGGTGGGTTGGG 588
QY	145 GCGAATTTACCGGTTGCTTTCTCTATCTTATTTCTCTCTCTCTCTCTCTCTGACC 204
Db	589 GGTAAATTGAGTGGCGGTTCTCTTTTCGGTTTCTTCGATTTCTCGGATTTCTTGAAA 648
QY	205 GTTCCGGCTCTGCAGTTCCTCCTACGAAATGCTCTGGATT 247
Db	649 GACAAAGAAGTGAAGGACCTTCATCAGATTTCTAGGCGGAATT 691
RESULT 8	
BI462364/c	
LOCUS	
DEFINITION	603203722F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269756 5', mRNA linear EST.
ACCESSION	BI462364
VERSION	GI:15253020
KEYWORDS	Genome
SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 721) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: cgabs@email.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRIL), Shiraki

Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
 Plate: LLAM11680 row: k column: 05
 High quality sequence stop: 712.
 Location/Qualifiers
 1. .721
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5269756"
 /lab_host="DH108"
 /clone_lib="NIH_MGC_97"
 /notes="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.2 kb and
 normalized to Rf 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIMH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 175 a 170 c 209 g 167 t

BASE COUNT
 ORIGIN

Query Match 5.8%; Score 36.6; DB 12; Length 721;
 Best Local Similarity 55.0%; Pred. No. 44;
 Matches 72; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 373 ACCCTTACTGTACGCCGAGCCCTCGGAGCAGTCAAGCTCTCTTCGAGAGCGGTT 432
 |||||
 Db 240 ACCACTAACTGGTTAGCCCCCATCTCTCCCGCGCGCGCTCGTCTCAAGCGCT 181
 |||||

QY 433 GACTACTAGCGGAGGCGTCCCTCTGTCGCGTTATACGTAGAGAGCGGTGGG 492
 |||||
 Db 180 CGTCTCCGCGGCGGAGCGGCTCTACCCCGCGGCGCTTCAGGGGCGCTTGAGCCGG 121
 |||||

QY 493 GCACTATTCTT 503
 |||||
 Db 120 GCACTAAGCGT 110
 |||||

RESULT 9
 CNS0140F/c
 LOCUS
 DEFINITION
 Drosophila melanogaster genome survey sequence 17 end of BAC
 BACN12U12 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL104409
 AL104409.1 GI:5616020
 GSS.
 Drosophila melanogaster (fruit fly)
 ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 962)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequençage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobAC11.

FEATURES
 source
 Location/Qualifiers
 1. .962
 /organism="Drosophila melanogaster"
 /mol_type="Genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACN12U12"
 /clone_lib="DrosBAC"
 /plasmid="pBelobAC11"
 /notes="end : T7"

BASE COUNT 226 a 156 c 169 g 207 t 204 others
 ORIGIN

Query Match 5.8%; Score 36.4; DB 29; Length 962;
 Best Local Similarity 35.3%; Pred. No. 52;
 Matches 110; Conservative 40; Mismatches 160; Indels 0; Gaps 0;

QY 231 AAATGCTCTGGATTATCATGTTACCAATGATGCCAAACTCTTCCATAGTCTATGA 290
 |||||
 Db 648 AATGGCGGCTTHACATDTTGTTCGCGGCACATCTCAWGTCTCCYAAARAAAAA 589
 |||||

QY 291 GGAGATAACCTGATCTACACGACCTGGTTCGTCGCTTGTGTGTCATGACAGTAATGT 350
 |||||
 Db 588 MAAKAAAAAATGAARCTCTGSCAATCACATTGSCATTTTCCCATGGRVRRKR 529
 |||||

QY 351 GAGTAGATGCTGGGTCAAATTAACCCCTACACTGTGAGCCCCAGCCTCGGAGCAGTCAC 410
 |||||
 Db 528 TRWCCGKKGATGGCACCAMVDMKCGGTAAACMSGATCAGCTGTCTACYAGCGGATCCSGCTY 469
 |||||

QY 411 GGCTCTCTTCGGAGAGCGGTTTGACTACCTAGCGGGGCGGCTCTGCTCCCGGTT 470
 |||||
 Db 468 TAGKCCGCCAGTGGAGGAGGAGKGYTKTGTGTTTGGAACAGGTATTGCTCGGGGAC 409
 |||||

QY 471 ATACGTAGGAGAGCGGTGTGGGCACACTATTTTGTGAGGCAAAATGTTACCTATAGGCC 530
 |||||
 Db 408 TGCNCGGKATGTCCKKBTGKCCCTCGTTTGTGCTCCACSCCCCYCNTTCCCC 349
 |||||

QY 531 TCGCCAGCAC 540
 |||||
 Db 348 CCCCHCCWC 339
 |||||

RESULT 10
 BX356664
 LOCUS
 DEFINITION
 BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CSOD1015YB03 3-PRIME, mRNA sequence.
 BX356664
 BX356664.1 GI:30378083
 EST.
 ACCESSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequençage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSOD1015CA02NP1.

FEATURES
 source
 Location/Qualifiers
 1. .1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1015YB03"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

sequence: 5'-ATTCTAGAGCCGAGGCGCCGACATG-dT(30)BN-3' (where B = A, C, G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

BASE COUNT	394 a	86 c	413 g	69 t
BASE COUNT	394 a	86 c	413 g	69 t

```
Query Match      5.6%; Score 35.4; DB 10; * Length 962;
Best Local Similarity 57.8%; Pred. No. 96;
Matches 63; Conservative 0; Mismatches 46; Indels 0
```

Qy	150	TTTTACGGGTGCTCTTTTCTCTAATCTTATCTTGCTCTCTCTCGGTGCTGACGGTTCC	209
Db	331	TTTCTCCTCTTCTTTC	272
Qy	210	GGCTCTGCAGTTCCTCTACCGAAATGCTCTGGGATTATCATGTTACC	258
Db	271	TCTTCTCTCTTTCGGCCCTTTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC	223

RESULT 14
BX397483/C

LOCUS	BX397483	1178 bp	mRNA	linear	EST 13-MAY-2003
DEFINITION	BX397483 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSDDI036YA06 5-PRIME, mRNA sequence.				

ACCESSION BX397483
VERSION , BX397483.1 GI:30631632
KEYWORDS EST.

RECORDS	SOURCE	ORGANISM	Homo sapiens (human)
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30
31	31	31	31
32	32	32	32
33	33	33	33
34	34	34	34
35	35	35	35
36	36	36	36
37	37	37	37
38	38	38	38
39	39	39	39
40	40	40	40
41	41	41	41
42	42	42	42
43	43	43	43
44	44	44	44
45	45	45	45
46	46	46	46
47	47	47	47
48	48	48	48
49	49	49	49
50	50	50	50
51	51	51	51
52	52	52	52
53	53	53	53
54	54	54	54
55	55	55	55
56	56	56	56
57	57	57	57
58	58	58	58
59	59	59	59
60	60	60	60
61	61	61	61
62	62	62	62
63	63	63	63
64	64	64	64
65	65	65	65
66	66	66	66
67	67	67	67
68	68	68	68
69	69	69	69
70	70	70	70
71	71	71	71
72	72	72	72
73	73	73	73
74	74	74	74
75	75	75	75
76	76	76	76
77	77	77	77
78	78	78	78
79	79	79	79
80	80	80	80
81	81	81	81
82	82	82	82
83	83	83	83
84	84	84	84
85	85	85	85
86	86	86	86
87	87	87	87
88	88	88	88
89	89	89	89
90	90	90	90
91	91	91	91
92	92	92	92
93	93	93	93
94	94	94	94
95	95	95	95
96	96	96	96
97	97	97	97
98	98	98	98
99	99	99	99
100	100	100	100

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4334.f For more information about this cluster, see

http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODI036BA030P1&cluster=434.f. Contact
Feng Liang Email: liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> InvitroGen Corporation 1600
Paradise Avenue Genoscope sequence ID: CSODI036BA030P1.

FEATURES	Location/Qualifiers
source	1. 1178

000000

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI036YA06"
/tissue_type="DIACENTA C

```

```

/culture_type=PLACENTA COL 25-NORMALIZED
/clone_lib=Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized.

```

BASE COUNT	267 a	118 c	175 g	236 t	382 others
ORIGIN					

Query Match 5.6%; Score 35.4; DB 13; Length 1178;
Best Local Similarity 19.3%; Pred. No. 99;
Matches 42; Conservative 85; Mismatches 91; Indels 0; Gaps 0

Qy	107	GTGTGAGGTCCTTGTAGGACGGGTTAACTATGCAACAGGGAATTACCGGTTGCTCTT	166
		: : : : : : : : : : : : : : : : : : : : : : : : : :	
Db	1104	DKKMAKINNDKKKKKKKGGTGTGMAASARKKGCMAWMDKAAABKKKKKKKKKKKKKKKK	1045

	Best Local Similarity	61.8%;	Pred.No.	1e+02;	Mismatches	34;	Indels	0;	Gaps	0;
	Matches	55;	Conservative	0;						
Qy	159	TGCGTCTTTTCATCATCTTTATTCTTGCTCCTCTCTCTGTGCTGTACGGTTCGGGCCTCTGCC	218							
Dd	512	TGGCAANTGGTTTTATAATAGCATGGCTTGCTTTTTTGTCTCTCCATTCAACTCTTCCC	571							
Qy	219	AGTTCCCTACGGAAATGCCCTCTGGGATTT	247							
Dd	572	TCTTCCTTACCATACTGCCCTGTGGGTTTT	600							

Search completed: December 20, 2003, 06:55:03
Job time : 1628.71.secs

Qy	167	TCTCTATCTTTATTCTGTCCTCTCGTGTCTGAACGGTTCGGGCCTCTGCAGTAGTCCCT	226
Db	1044	TTACKTCTCTBKKKBAAKTKCKTKMBAAMATKKMBMTKKTKTKTCCCCTKNTBACCMTK	985
Qy	227	ACGGAAATGCCCTCTGGGATTTATCATGTTACAATGATGCCCAAACACTCTTCCATAGTCT	286
Db	984	BMCCYKKKCKCWMHCYCWTKTTMCCTTMMCTKTNTTMCMVMCMCTTKTYCKMKNCMT	925
Qy	287	ATGAGGAGATAACCTTGATCTCAACGACCCTGGTTGC	324
bh	924	HTKMCCYGMGMHMCCKCCBVKTKKMMCKCCOMMKKK	887

[illegible]

FEATURES	source
Location/Qualifiers	
1	27
	/organism="Gallus gallus"
	/moi_type="mRNA"
	/strain="Layer and broiler"
	/db_xref="taxon:9031"
	/clone="CHEST938d"
	/sex="Male and female"
	/tissue_type="Chondrocytes isolated from growth plate cartilage"
	/dev_stage="adult"
	/lab_host="DH10B"
	/lab_host="DH10B"

/clone_lib="L5696903
/note="vector: pBluescript II KS(+); Site_1: EcoRI;
Site_2: NotI; This normalized library was constructed from
1 million independent clones. cDNA synthesis was initiated
using an oligo(dT) primer, using methylated C in the first
strand synthesis reaction. Following this first strand
reaction, double-stranded cDNA was blunted, ligated to
NotI adapters, digested with EcoRI, size-selected, and
cloned into the NotI and EcoRI compatible sites of a
custom modified MCS of the pBluescript (KS+) vector. The
library was normalized in 2 rounds using conditions
adapted from Soares et al., PNAS (1994) 91: 9282-9322 and
Bonaldi et al., Genome Research 6 (1996): 791, except that
a significantly longer reannealing hybridization was
used."

BASE COUNT	207 a	122 c	149 g	248 t	1 others
used:					

Query Match	Score	DB 13;	DB 13;	Length 727;
	5.6%	Score 35.2;	DB 13;	Length 727;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 45.4354 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-31
Perfect score: 630
Sequence: 1 ATGGGTAGTCATCGATAC.....TGATGATGAACCTGGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	100.0	630	3	US-08-612-973-31
2	630	100.0	630	3	US-08-927-597-31
3	347.2	55.1	576	1	US-08-086-428B-47
4	347.2	55.1	576	2	US-08-468-570-47
5	347.2	55.1	576	2	US-08-290-665A-47
6	347.2	55.1	576	4	US-08-466-601A-47
7	347.2	55.1	576	5	PCT-US95-10398-47
8	344	54.6	576	1	US-08-086-428B-49
9	344	54.6	576	2	US-08-468-570-49
10	344	54.6	576	2	US-08-290-665A-49
11	344	54.6	576	4	US-08-466-601A-49
12	344	54.6	576	5	PCT-US95-10398-49
13	340.8	54.1	576	1	US-08-086-428B-48
14	340.8	54.1	576	2	US-08-468-570-48
15	340.8	54.1	576	2	US-08-290-665A-48
16	340.8	54.1	576	4	US-08-466-601A-48
17	340.8	54.1	576	5	PCT-US95-10398-48
18	337.6	53.6	576	1	US-08-086-428B-46
19	337.6	53.6	576	2	US-08-468-570-46
20	337.6	53.6	576	2	US-08-290-665A-46
21	337.6	53.6	576	4	US-08-466-601A-46
22	337.6	53.6	576	5	PCT-US95-10398-46
23	337.6	53.6	576	2	US-08-290-665A-50
24	337.6	53.6	576	4	US-08-466-601A-46
25	337.6	53.6	576	4	US-08-466-601A-50
26	337.6	53.6	576	5	PCT-US95-10398-46
27	337.6	53.6	576	5	PCT-US95-10398-50

28	335.6	53.3	630	3	US-08-612-973-29	Sequence 29, Appl
29	335.6	53.3	630	3	US-08-927-597-29	Sequence 29, Appl
30	331.6	52.6	9595	3	US-09-014-416-4	Sequence 4, Appl
31	331.6	52.6	9599	3	US-09-014-416-6	Sequence 6, Appl
32	331.2	52.6	576	1	US-08-086-428B-45	Sequence 45, Appl
33	331.2	52.6	576	2	US-08-468-570-45	Sequence 45, Appl
34	331.2	52.6	576	2	US-08-290-665A-45	Sequence 45, Appl
35	331.2	52.6	576	4	US-08-466-601A-45	Sequence 45, Appl
36	331.2	52.6	576	5	PCT-US95-10398-45	Sequence 45, Appl
37	325.2	51.6	1539	2	US-08-470-426B-17	Sequence 17, Appl
38	325.2	51.6	1863	2	US-08-470-426B-14	Sequence 14, Appl
39	324.8	51.6	795	3	US-08-612-973-5	Sequence 5, Appl
40	324.8	51.6	795	3	US-08-927-597-5	Sequence 5, Appl
41	324.8	51.6	2082	3	US-08-612-973-47	Sequence 47, Appl
42	324.8	51.6	2082	3	US-08-927-597-47	Sequence 47, Appl
43	324.8	51.6	2433	3	US-08-612-973-49	Sequence 49, Appl
44	324.8	51.6	2433	3	US-08-927-597-49	Sequence 49, Appl
45	324.8	51.6	9472	4	US-08-150-204E-96	Sequence 96, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-31
; Sequence 31, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..627
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..624


```

US-08-612-973-31

Query Match          100.0%; Score 630; DB 3; Length 630;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGGTAAGGTGATCATGATACCTTAACGTCGGGATTCGCCGATCTCATGGGGTATATCCCG 60
1 ATGGGTAAGGTGATCATGATACCTTAACGTCGGGATTCGCCGATCTCATGGGGTATATCCCG 60
61 CTCGTAGGGGGCCCATTTGGGGGCGTCGACAGGGCTCTCGCACACGGTGTGAGGGTCCTT 120
61 CTCGTAGGGGGCCCATTTGGGGGCGTCGACAGGGCTCTCGCACACGGTGTGAGGGTCCTT 120
121 GAGGACGGGTAACTATCAACAGGGAATTTACCCGGTTCCTCTTCTCTATCTTTATT 180
121 GAGGACGGGTAACTATCAACAGGGAATTTACCCGGTTCCTCTTCTCTATCTTTATT 180
181 CTTGCTCTTCTCGTGTCTGACCGTTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT 240
181 CTTGCTCTTCTCGTGTCTGACCGTTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT 240
241 GGGATTTATCATGTTTACCAATGATGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT 300
241 GGGATTTATCATGTTTACCAATGATGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT 300
301 CTGATCCTACAGCACCTGTTGTCGTCTGACCGTTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT 360
301 CTGATCCTACAGCACCTGTTGTCGTCTGACCGTTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT 360
361 TGGGTCCAAATTTACCCCTACACTGTGACCGTTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT 420
361 TGGGTCCAAATTTACCCCTACACTGTGACCGTTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT 420
421 CGGAGACGGGTGAATCACTAGCGGAGGGGTCGCTCTGTCGCGGTTATAGTAGGA 480
421 CGGAGACGGGTGAATCACTAGCGGAGGGGTCGCTCTGTCGCGGTTATAGTAGGA 480
481 GACCGGTGGGGCACTATTTCTGGTAGGCAAAATGTTTACAGTGGCCATGTTACCGGCCACGGATG 540
481 GACCGGTGGGGCACTATTTCTGGTAGGCAAAATGTTTACAGTGGCCATGTTACCGGCCACGGATG 540
541 GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACGGATG 600
541 GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACGGATG 600
601 GCATGGGATGATGATGAACTGGTAATAG 630
601 GCATGGGATGATGATGAACTGGTAATAG 630

RESULT 2
US-08-927-597-31
; Sequence 31, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P. C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```


QY 220 GTTCCCTACCGAATCCCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 279
Db 1 GTCCCTTACCGAATCCCTCTGGGTTTATCATGTACCAATGATGCCCAAACTCTTCC 60
QY 280 ATAGTCTATGAGGACATACCTGATCTTACGACGACCTGTTGGTGCTGTGTCATG 339
Db 61 ATAGTCTACGAGGCTGATACCTGATCTTGCAGCACCTGTTGGTGCTGTGTCATG 120
QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATTTACCCCTACACTGTTCAGCCCGGAGCCTC 399
Db 121 GAAGGTAATGTAGTAGGCTGGGTCCAAATTTACCCCTACACTGTTCAGCCCGGAGCCTC 180
QY 400 GGAGCAGTACAGGCTCTCTTTCGGAGAGCGGTTGACTACCTAGCGGAGGAGGCTGCCCTC 459
Db 181 GGAGCGGTACAGGCTCTCTTTCGGAGAGCGGTTGACTACCTAGCGGAGGAGGCTGCCCTC 240
QY 460 TGCTCCGCTTATACGTAGGAGCGGTGGGGCACTATTCTGGTAGGCCAAATGTTTC 519
Db 241 TGCTCCGCACTATACGTGGGAGCGGTGGGGCACTATTCTGGTAGGCCAAATGTTTC 300
QY 520 ACCTATAGGCTCGCCAGCACGCTACGGTGCAGAACTGCAACTGTTTCCATTTTACAGTGGC 579
Db 301 ACCTATAGGCTCGCCAGCATACTACGGTGCAGAACTGCAACTGTTTCCATTTTACAGCGGC 360
QY 580 CATGTTACCGCCACCGATGGCATGGGATATGATGATGAATGGTAA 627
Db 361 CATATCACCGCCACCGAATGGCATGGGACATGATGATGAATGGTCA 408

RESULT 5

US-08-290-665A-47
; Sequence 47, Application US/08290665A

; Patent No. 5882852

; GENERAL INFORMATION:

; APPLICANT: BUKH, J., MILLER, R.H. AND

; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED

; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND

; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS

; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE

; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES

; NUMBER OF SEQUENCES: 263

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/290,665A

; FILING DATE: 15-AUG-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: RICHARD W. BORK

; REGISTRATION NUMBER: 36,459

; REFERENCE/DOCKET NUMBER: 2026-4116

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 576 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; ORIGINAL SOURCE:

; ORGANISM: hom sapiens
; INDIVIDUAL ISOLATE: SAS
US-08-290-665A-47

Query Match 55.1%; Score 347.2; DB 2; Length 576;

Best Local Similarity 90.7%; Pred. No. 1.5e-101;

Matches 370; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAATCCCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 279

Db 1 GTCCCTTACCGAATCCCTCTGGGTTTATCATGTACCAATGATGCCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGGACATACCTGATCTTACGACGACCTGTTGGTGCTGTGTCATG 339

Db 61 ATAGTCTACGAGGCTGATACCTGATCTTGCAGCACCTGTTGGTGCTGTGTCATG 120

QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATTTACCCCTACACTGTTCAGCCCGGAGCCTC 399

Db 121 GAAGGTAATGTAGTAGGCTGGGTCCAAATTTACCCCTACACTGTTCAGCCCGGAGCCTC 180

QY 400 GGAGCAGTACAGGCTCTCTTTCGGAGAGCGGTTGACTACCTAGCGGAGGAGGCTGCCCTC 459

Db 181 GGAGCGGTACAGGCTCTCTTTCGGAGAGCGGTTGACTACCTAGCGGAGGAGGCTGCCCTC 240

QY 460 TGCTCCGCTTATACGTAGGAGCGGTGGGGCACTATTCTGGTAGGCCAAATGTTTC 519

Db 241 TGCTCCGCACTATACGTGGGAGCGGTGGGGCACTATTCTGGTAGGCCAAATGTTTC 300

QY 520 ACCTATAGGCTCGCCAGCACGCTACGGTGCAGAACTGCAACTGTTTCCATTTTACAGTGGC 579

Db 301 ACCTATAGGCTCGCCAGCATACTACGGTGCAGAACTGCAACTGTTTCCATTTTACAGCGGC 360

QY 580 CATGTTACCGCCACCGATGGCATGGGATATGATGATGAATGGTAA 627

Db 361 CATATCACCGCCACCGAATGGCATGGGACATGATGATGAATGGTCA 408

RESULT 6

US-08-466-601A-47

; Sequence 47, Application US/08466601A

; Patent No. 6572864

; GENERAL INFORMATION:

; APPLICANT: BUKH, J., MILLER, R.H. AND

; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED

; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE

; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE

; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN

; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES

; NUMBER OF SEQUENCES: 160

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/466,601A

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA: 08/086,428

; APPLICATION NUMBER: 08/086,428

; FILING DATE: 29-JUN-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: RICHARD W. BORK

; REGISTRATION NUMBER: 36,459


```
; REFERENCE/DOCKET NUMBER: 2026-4070US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA5
; US-08-466-601A-47

Query Match 55.1%; Score 347.2; DB 4; Length 576;
Best Local Similarity 90.7%; Pred. No. 1.5e-101;
Matches 370; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAAATGCTGGGATTATCATGTTACCAATGATGCCCCAAACTCTTCC 279
Db |||||
1 GTCCCTACCGAAATGCTGGGTTTATCATGTACCAATGATGCCCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGCAGATACCTGATCCCTACAGCACCTGGTGGCGTCTGTGTCATG 339
Db |||||
61 ATAGTCTACGAGGCTGATACCTGATCTGCAGCACCTGGTGGCGTCTGTGTCATG 120

QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATACCCCTACACTGTGAGCCCGAGCCTC 399
Db |||||
121 GAAGGTAATGTAGTAGGCTGGGTCCAAATACCCCTACACTGTGAGCCCGAGCCTC 180

QY 400 GGAGCAGTCACGGCTCTCTTCGGAGAGCGGTGAGTACCTAGCGGAGGGGTGCGCCTC 459
Db |||||
181 GGAGCGGTACGGCTCTCTTCGGAGAGCGGTGAGTACCTAGCGGAGGGGTGCGCCTC 240

QY 460 TGCTCCGCTTATACGTAGAGACGCGTGTGGGGCACTATTTGGTAGCCAAATGTTTC 519
Db |||||
241 TGCTCCGCACTATACGTGCGGGACGCGTGGGGCAGTGTCTTGGTAGCCAAATGTTTC 300

QY 520 ACCTATAGGCTCGCCAGCACGCTACGCTGCAGAACTGCAACTGTTCCATTACAGTGC 579
Db |||||
301 ACCTATAGGCTCGCCAGCATATACGCTGCAGAACTGCAACTGTTCCATTACAGGCGC 360

QY 580 CATGTTACCGGCCACCGATGGCATGGGATATGATGATGAATGTTAA 627
Db |||||
361 CATATCAGCGCCACCGATGGCATGGGATATGATGATGAATGTTAA 408

RESULT 7
PCT-US95-10398-47
; Sequence 47, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA5
; PCT-US95-10398-47

Query Match 55.1%; Score 347.2; DB 5; Length 576;
Best Local Similarity 90.7%; Pred. No. 1.5e-101;
Matches 370; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAAATGCTGGGATTATCATGTTACCAATGATGCCCCAAACTCTTCC 279
Db |||||
1 GTCCCTACCGAAATGCTGGGTTTATCATGTACCAATGATGCCCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGCAGATACCTGATCCCTACAGCACCTGGTGGCGTCTGTGTCATG 339
Db |||||
61 ATAGTCTACGAGGCTGATACCTGATCTGCAGCACCTGGTGGCGTCTGTGTCATG 120

QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATACCCCTACACTGTGAGCCCGAGCCTC 399
Db |||||
121 GAAGGTAATGTAGTAGGCTGGGTCCAAATACCCCTACACTGTGAGCCCGAGCCTC 180

QY 400 GGAGCAGTCACGGCTCTCTTCGGAGAGCGGTGAGTACCTAGCGGAGGGGTGCGCCTC 459
Db |||||
181 GGAGCGGTACGGCTCTCTTCGGAGAGCGGTGAGTACCTAGCGGAGGGGTGCGCCTC 240

QY 460 TGCTCCGCTTATACGTAGAGACGCGTGTGGGGCACTATTTGGTAGCCAAATGTTTC 519
Db |||||
241 TGCTCCGCACTATACGTGCGGGACGCGTGGGGCAGTGTCTTGGTAGCCAAATGTTTC 300

QY 520 ACCTATAGGCTCGCCAGCACGCTACGCTGCAGAACTGCAACTGTTCCATTACAGTGC 579
Db |||||
301 ACCTATAGGCTCGCCAGCATATACGCTGCAGAACTGCAACTGTTCCATTACAGGCGC 360

QY 580 CATGTTACCGGCCACCGATGGCATGGGATATGATGATGAATGTTAA 627
Db |||||
361 CATATCAGCGCCACCGATGGCATGGGATATGATGATGAATGTTAA 408

RESULT 8
US-08-086-428B-49
; Sequence 49, Application US/08086428B
; Patent No. 5514539
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
```



```
;
;
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
;
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/086,428B
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6849
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA7
;
; US-08-086-428B-49
;
; Query Match 54.6%; Score 344; DB 1; Length 576;
; Best Local Similarity 90.2%; Pred. No. 1.5e-100;
; Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
;
; QY 220 GTTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATTCGCCAAACTCTTCC 279
; DB 1 GTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATTCGCCAAACTCTTCC 60
;
; QY 280 ATAGTCTATGAGGAGATAACCTGATCTACACGACCTGTTGCGTGTGTCATG 339
; DB 61 ATAGTCTATGAGGAGCTGACCACTGATCTTCGACGACCTGTTGCGTGTGTCAGA 120
;
; QY 340 ACAGGTAATGTAGTAGTGTGGTCCAAATACCCCTACACTGTACGCCCCGAGCCTC 399
; DB 121 CAAATAATGTAGTAGTGTGGTCCAAATACCCCTACACTGTACGCCCCGAGCCTC 180
;
; QY 400 GGAGCAGTCAGGCTCTCTTCGAGAGCCGTGACTACCTAGCGGGAGGGGCTGCCCTC 459
; DB 181 GGAGCGGTACGGCTCTCTTCGAGAGCCGTGACTACCTAGCGGGAGGGGCTGCCCTC 240
;
; QY 460 TGCTCCGGTTATACGTAGGAGACCGGTGTGGGGCACTATTCTTGGTAGGCCAAATGTT 519
; DB 241 TGCTCCGGCTATACGTGGGAGACCGGTGTGGGGCAGTGTCTTGGTAGGCCAGATGTT 300
;
; QY 520 ACCTATAGGCTCGCAGCAGCTACGCTGCGAGACTGCACTGTTCCATTTACAGTGGC 579
; DB 301 AGCTATAGGCTCGCAGCAGCTACGCTGCGAGACTGCACTGTTCCATTTACAGTGGC 360
;
; QY 580 CATGTTACCGGCACCGGATGGATGGATATGATGATGAATGTA 627
; DB 361 CATATACCGGCACCGGATGGATGGATGATGATGAATGTA 408
;
;
; RESULT 9
; US-08-468-570-49
; Sequence 49, Application US/08468570
;
; Patent No. 5871962
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,570
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6849
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA7
;
; US-08-468-570-49
;
; Query Match 54.6%; Score 344; DB 2; Length 576;
; Best Local Similarity 90.2%; Pred. No. 1.5e-100;
; Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
;
; QY 220 GTTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATTCGCCAAACTCTTCC 279
; DB 1 GTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATTCGCCAAACTCTTCC 60
;
; QY 280 ATAGTCTATGAGGAGATAACCTGATCTACACGACCTGTTGCGTGTGTCATG 339
; DB 61 ATAGTCTATGAGGAGCTGACCACTGATCTTCGACGACCTGTTGCGTGTGTCAGA 120
;
; QY 340 ACAGGTAATGTAGTAGTGTGGTCCAAATACCCCTACACTGTACGCCCCGAGCCTC 399
; DB 121 CAAATAATGTAGTAGTGTGGTCCAAATACCCCTACACTGTACGCCCCGAGCCTC 180
;
; QY 400 GGAGCAGTCAGGCTCTCTTCGAGAGCCGTGACTACCTAGCGGGAGGGGCTGCCCTC 459
; DB 181 GGAGCGGTACGGCTCTCTTCGAGAGCCGTGACTACCTAGCGGGAGGGGCTGCCCTC 240
;
; QY 460 TGCTCCGGTTATACGTAGGAGACCGGTGTGGGGCACTATTCTTGGTAGGCCAAATGTT 519
; DB 241 TGCTCCGGCTATACGTGGGAGACCGGTGTGGGGCAGTGTCTTGGTAGGCCAGATGTT 300
;
; QY 520 ACCTATAGGCTCGCAGCAGCTACGCTGCGAGACTGCACTGTTCCATTTACAGTGGC 579
; DB 301 AGCTATAGGCTCGCAGCAGCTACGCTGCGAGACTGCACTGTTCCATTTACAGTGGC 360
;
; QY 580 CATGTTACCGGCACCGGATGGATGGATATGATGATGAATGTA 627
; DB 361 CATATACCGGCACCGGATGGATGGATGATGATGAATGTA 408
;
;
; RESULT 9
; US-08-468-570-49
; Sequence 49, Application US/08468570
;
```


220	GTTCCTACCGAAATGCCTCTGGGATTTATCATGTTACCAATGATTGCCAACTCTTCC	279
Db	1	GTCCCTACCGNAATGCCTCCGGGTTTATCATGTACCAATGATTGCCGAACCTCTCC
280	ATAGTCTATGAGGCAGATAA	339
Db	61	ATAGTCTATGAGGCTGACAACGTATCTGTGACGCACCTGTTGCGTGCCTGTGTGAGA
340	ACAGGTATGTGAGTAGATGCTGGTCCAAATTACCCCTACACTGT	399
Db	121	CAAAATAATGT
400	GGAGCAGTCA	459
Db	181	GGAGCGGTCA
460	TGCTCGCGTTATAGCTAGGAGACCGGTGGGGCACTATCTTGTTGATGGCCAAATGTTTC	519
Db	241	TGCTCGCGCTATACGTGCGGACCGGTGCGGGGCGAGTGTGTTGTTGATGGCCAGATGTTTC
520	ACCTATAGGCGCTCGCCAGCACGCTACGGTGGAGAACTGCAACTGTTCCATTTACAGTGGC	579
Db	301	AGCTATAGCGCTCGCCAGCACACTACGGTGCAGGACTGCAACTGTTCCATTTACAGTGGC
580	CATGTTACCGGCCACCGGATGGCATGGAATGATGAACTGGTAA	627
Db	361	CATATCACCGGCCACCGAATGGCATGGGACATGATGAAATGGTCA

RESULT 12

PCT-US95-10398-49
; Sequence 49, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BURKH, J., MILLER, R.H. AND
; APPLICANT: BURKH, J., MILLER, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:

TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES

;/
;/ NUMBER OF SEQUENCES: 263
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: MORGAN & FINNEGAN
;/ STREET: 345 PARK AVENUE
;/ CITY: NEW YORK
;/ STATE: NEW YORK
;/ COUNTRY: USA
;/ ZIP: 10154
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: FLOPPY DISK
;/ COMPUTER: IBM PC COMPATIBLE
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: WORDPERFECT 5.1
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/290,665A
;/ FILING DATE: 15-AUG-1994
;/ CLASSIFICATION: 435
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: RICHARD W. BORK
;/ REGISTRATION NUMBER: 36,459
;/ REFERENCE/DOCKET NUMBER: 2026-4116
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (212) 751-4800
;/ TELEFAX: (212) 751-6849
;/ TELEX: 421792
;/ INFORMATION FOR SEQ ID NO: 48:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 576 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ ORIGINAL SOURCE:
;/ ORGANISM: homosapiens
;/ INDIVIDUAL ISOLATE: SA6
;/ US-08-290-665A-48

Query Match 54.1%; Score 340.8; DB 2; Length 576;
Best Local Similarity 89.7%; Pred. No. 1.6e-99;
Matches 366; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 220 GTTCCTACCGAATGCCTCTGGGATTTCATGTTACCAATGATTCGCCAAACTCTTCC 279
Db 1 GTTCCTACCGAATGCCTCTGGGATTTCATGTTACCAATGATTCGCCAAACTCTTCC 60

QY 280 ATAGTCTATAGGCGAGATAACCTGATCCTACACGACCTGGTTCGTCCTTGTGTGTCATG 339
Db 61 ATAGTCTATAGGCGAGATAACCTGATCCTACACGACCTGGTTCGTCCTTGTGTGTCATG 120

QY 340 ACAGGTAATGTGAGTAGATGCTGGTCCAAATTCACCTACACTGTGAGCCCGAGCCTC 399
Db 121 AAGGTAATGTGAGTAGATGCTGGTCCAAATTCACCTACACTGTGAGCCCGAGCCTC 180

QY 400 GGAGCAGTTCACGGCTCCTCTTCGGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTC 459
Db 181 GGAGGCTTCACGGCTCCTCTTCGGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTC 240

QY 460 TGCTCCGGGTTATACGTAGAGAGCGCGTGTGGGCACTATCTTGTGAGCCAAATGTTTC 519
Db 241 TGCTCCGGGTTATACGTAGAGAGCGCGTGTGGGCACTATCTTGTGAGCCAAATGTTTC 300

QY 520 ACCTATAGGCTCCGACGACGCTACCGTGCAGACTGCAACTGTTCCATTACAGTGGC 579
Db 301 ACCTATAGGCTCCGACGACGCTACCGTGCAGACTGCAACTGTTCCATTACAGTGGC 360

QY 580 CATGTTACCGGCCACCGGATGGCATGGGATATGATGATGAACCTGGTAA 627
Db 361 CATATCACTGGCCACCGGATGGCATGGGATATGATGATGAACCTGGTAA 408